

UNIVERSITÉ DE MONTRÉAL

ÉTUDE D'UN FILTRE NEURONAL BASÉ SUR LA STRUCTURE DU  
FILTRE DE KALMAN

LÉANDRE NNEME NNEME  
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE ET DE GÉNIE INFORMATIQUE  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION DU DIPLÔME DE  
PHILOSOPHIAE DOCTOR (Ph.D.)  
(GÉNIE ÉLECTRIQUE)  
MARS 1998



National Library  
of Canada

Acquisitions and  
Bibliographic Services

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque nationale  
du Canada

Acquisitions et  
services bibliographiques

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-33014-1

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

ÉTUDE D'UN FILTRE NEURONAL BASÉ SUR LA STRUCTURE DU  
FILTRE DE KALMAN

Présentée par: NNEME NNEME Léandre

en vue de l'obtention du diplôme de: Philosophiae Doctor

a été dûment acceptée par le jury d'examen constitué de:

M. HURTEAU Richard D. Ing, Président

M. O'SHEA Jules D.Ing., membre et directeur de recherche

M. NERGUIZIAN Chahé M. Ing., membre et codirecteur de recherche

M. SAAD Maarouf Ph.D., membre

M. BRAULT Jean Jules Ph.D., membre

# Dédicace

A mon épouse, Martine

A mes chers enfants, Oswald Jordane et Karl Kevin

A mon feu papa, Pierre Nneme Edou

A ma mère, Grégorienne Mfono Emvo

A mes frères et soeurs

Au feu, docteur Jean Assoumou Avebe

# Remerciements

Qu'il me soit permis d'exprimer ma très grande reconnaissance d'abord à Monsieur Jacques Martin (coordonateur des boursiers de l'ACDI à l'École Polytechnique) et à travers lui l'ACDI (Agence Canadienne de Développement International) et ensuite à l'Université de Douala qui m'ont donné l'opportunité de me libérer afin de poursuivre mes études doctorales.

Je tiens à exprimer mes très sincères remerciements à mon directeur de recherche le Professeur Jules O'Shea pour son esprit scientifique très riche dont il a su m'inculquer la substance tout au long de cette thèse. Il n'a ménagé aucun effort pour m'apporter, en tout temps et en tout lieu, l'encadrement nécessaire et appréciable qui m'a mis à l'abri des difficultés certaines. Notre interaction a été parfaite. Sa patience, sa coopération, son sens très poussé de responsabilité, son humilité et sa simplicité ont créé entre nous un réel climat de complicité et d'entente. J'adresse mes remerciements à mon codirecteur M. Chahé Nerguizian pour ses conseils et ses commentaires pertinents.

Mes remerciements vont également aux membres du jury, M. Richard Hurteau, M. Jean Jules Brault, M. Robert LeBlanc A. et particulièrement M. Maarouf Saad pour avoir accepté d'évaluer cette thèse.

J'aimerais remercier tous les professeurs de la section automatique pour leurs conseils salvateurs tout au long de ce travail et particulièrement le professeur Roland

P. Malhamé pour ses explications et ses suggestions pertinentes.

Qu'il me soit, en outre, permis de remercier très profondément M. John Thistle pour la confiance qu'il m'a accordée tout au long de mon séjour à l'École Polytechnique en me donnant diverses responsabilités d'enseignement.

Je remercie M.Yves Léonard et M. Richard Grenier, techniciens à la section automatique, pour leur dévouement à nous fournir en tout temps les outils nécessaires pour l'avancement de nos activités d'enseignement et de recherche.

A tous mes camarades de la section automatique, je dis merci pour la chaleur humaine que vous avez su maintenir autour de moi pendant mon séjour dans la section.

Je remercie enfin mes amis le docteur Gerba Malam et M. Daniel Roger Ngoa Nguele, entre autres, pour avoir accepté de revoir la forme de cette thèse.

## Résumé

Le besoin constant de connaître le futur se fait le plus souvent ressentir en commande où la nécessité d'avoir l'information sur le futur est primordiale dans le but de calculer la commande adéquate. La commande prédictive est l'un des moyens utilisés par les chercheurs pour atteindre cet objectif. Dans l'état actuel des connaissances, le filtre de Kalman est le meilleur estimateur linéaire à cause de l'optimalité de son gain. Il est néanmoins limité dans sa formulation; d'une part, la prédiction qu'il fait est juste à un pas dans le temps, d'autre part, le gain optimal est une fonction des statistiques des bruits de modèle et de mesure du système sans lesquelles son implantation est quasi impossible et l'optimalité du gain n'est plus garantie. Le présent travail cherche à remplacer le gain du filtre de Kalman, dans sa structure, par un réseau de neurones artificiels (RNA). La nouvelle structure est appelée *le filtre neuronal*. Suivant le cas, ce filtre peut être un estimateur ou un prédicteur. La comparaison des résultats obtenus par cette nouvelle configuration à ceux donnés par le filtre de Kalman, dans l'hypothèse d'observations connues et incomplètes, montre que le filtre neuronal est un bon estimateur et un bon prédicteur des états des systèmes dynamiques et qu'il s'adapte le mieux aux variations des paramètres des systèmes. Il faut cependant noter que la qualité des résultats produits par le filtre neuronal est dépendante du degré de bruit du système dont on veut prédire ou estimer les états. Si le système est très corrompu, les résultats donnés par le

filtre se détériorent. A partir des informations entrée/sortie du système dont on dispose, le RNA procède à un apprentissage au bout duquel il acquerra la capacité de prédire les états du système à des instants donnés. Il sera ainsi possible de prédire le comportement du système assez loin dans le futur c'est-à-dire construire un prédicteur neuronal à long terme. Une application pratique sur les réservoirs hydrologiques d'Hydro-Québec (Baskatong, Cabonga, Maniwaki) permet de valider cet aspect prédictif du filtre neuronal. Le filtre neuronal et le filtre de Kalman, dans le contexte d'informations connues, aboutissent presque aux mêmes résultats. Si les informations sont partielles, sans toutefois faire une prédiction parfaite, le prédicteur neuronal s'avère être mieux indiqué pour la prédiction. Par ailleurs, il n'est pas nécessaire de connaître les statistiques des bruits lors de cette prédiction; ce qui n'est pas le cas du filtre de Kalman. Une autre application sur l'estimation des états d'un avion Boeing 747 en commande du mouvement longitudinal montre que le filtre neuronal fait une bonne estimation des variables d'état du système en étude. Si les statistiques des bruits et le modèle du système sont parfaitement connus, le filtre de Kalman demeure cependant le meilleur estimateur linéaire. Lorsque la matrice d'état du système est perturbée, les résultats de simulation montrent que les estimés donnés par le filtre de Kalman se détériorent alors que l'estimateur neuronal est quasi insensible aux variations des paramètres du modèle jusqu'à un rapport signal-bruit donné. Dans un tel contexte, il est préférable d'utiliser le filtre neuronal à cause de sa presque insensibilité aux variations des paramètres du système.



# Abstract

Man has always sought to improve his vision of the future by more or less successfully devising means for getting information on events to come. The need for determining ahead of time the course of events becomes more crucial when one has to formulate the control law for a system with delay. The predictive control is one of the type of control used by researchers to accomplish this need; but actually Kalman filter constitutes the best state of the art linear estimator due to the optimality of its gain matrix. However, its major limitation comes from the fact that firstly, the Kalman filter yields only a one-step ahead estimation, secondly, its optimal matrix gain is a function of the model noise and the measurement noise whose ignorance makes the implementation of the filter practically impossible and the optimality guarantee of the matrix gain falls. This work bears precisely on the use of Artificial Neural Network (ANN) for building long term estimators in noisy environment. The gain matrix of Kalman filter is replaced by an ANN and the results of the neural filter are compared with those of Kalman filter in the context of either complete information or incomplete information. From the input/output data of the system which are available, the ANN goes into a learning process that eventually gives it the capability to estimate the state variables of the dynamical system at discrete instants. It will be seen that this will permit to predict many step ahead into the future thus offering the possibility to build a long term neural

predictor which is insensitive to system parameters variation when the signal to noise ratio is small.

One application is carried out to predict from archival data the level of water in a system made up of three reservoirs in cascade which belong to Hydro-Québec and are named Baskatong, Maniwaki, Cabonga. Simulation results of both the Kalman filter and the neural network allow one to compare the performances of these two predicting structures. Both follow quite well the actual values. In case of incomplete or missing informations, the Kalman filter can no longer follow the actual values while the neural network can still predict correctly some values if the signal to noise ratio is small. Moreover, one does not need to know the statistics of the noises for the neural predictor which is not so for the Kalman filter.

One more example concerns the estimation of the state variables of a Boeing 747 aircraft. The simulation of its longitudinal motion shows that the neural estimator gives good estimations of the state variables. But if the state matrix of the system is perturbed, the Kalman filter gives the worst estimates. It remains that the Kalman filter is superior when the noise statistics and the model parameters are known. Then, one can benefit by using the neural filter since it is much less affected by the parameters variations.

# Table des matières

<b>Dédicace .....</b>	<b>iv</b>
<b>Remerciements .....</b>	<b>v</b>
<b>Résumé .....</b>	<b>vii</b>
<b>Abstract .....</b>	<b>ix</b>
<b>Table des matières .....</b>	<b>xi</b>
<b>Liste des figures .....</b>	<b>xix</b>
<b>Liste des tableaux .....</b>	<b>xxvii</b>
<b>Liste des symboles .....</b>	<b>xxix</b>
<b>Liste des annexes .....</b>	<b>xxxii</b>
<b>Chapitre 1:        INTRODUCTION GÉNÉRALE .....</b>	<b>1</b>

<b>Chapitre 2:</b>	<b>MÉTHODES DE DÉTERMINATION</b>	
	<b>DU GAIN DU FILTRE DE KALMAN . . . . .</b>	<b>7</b>
2.1	GÉNÉRALITÉS SUR LE FILTRE DE KALMAN . . . . .	7
2.1.1	Le filtre de Kalman étendu . . . . .	12
2.1.1.1	Identification des paramètres par le filtre de Kalman	13
2.1.2	Stabilité et Convergence du Filtre de Kalman . . . . .	14
2.2	QUELQUES MÉTHODES DE DÉTERMINATION DU GAIN DU FILTRE DE KALMAN . . . . .	17
2.2.1	Méthodes indirectes . . . . .	17
2.2.1.1	Méthode de corrélation de Koji Tajima . . . . .	17
2.2.1.2	Méthode de Bergman et Delleur . . . . .	19
2.2.1.3	Méthode de Assaf et Quick . . . . .	21
2.2.1.4	Méthode de Todini . . . . .	21
2.2.1.5	Méthode de Mayers et Tapkey . . . . .	21
2.2.2	Méthodes directes . . . . .	22
<b>Chapitre 3:</b>	<b>NOTIONS DE BASE SUR LES</b>	
	<b>RÉSEAUX NEURONAUX . . . . .</b>	<b>28</b>

3.1	FONDEMENTS BIOLOGIQUES . . . . .	28
3.1.1	Dendrites . . . . .	29
3.1.2	Le corps cellulaire . . . . .	29
3.1.3	Synapses . . . . .	29
3.1.4	L'axone . . . . .	30
3.1.5	Fonctionnement des neurones biologiques . . . . .	30
3.2	FONDEMENT SUR LES NEURONES ARTIFICIELS . . . . .	32
3.2.1	Représentation . . . . .	33
3.2.2	Topologie . . . . .	35
3.2.2.1	Réseaux à couches . . . . .	35
3.2.2.2	Réseaux entièrement connectés . . . . .	36
3.2.3	Propagation des signaux . . . . .	37
3.3	DIFFÉRENTS MODÈLES DE RÉSEAUX NEURONAUX ARTIFI- CIELS . . . . .	37
3.4	APPRENTISSAGE DES RÉSEAUX . . . . .	38
3.4.1	Apprentissage supervisé . . . . .	39
3.4.2	Apprentissage non supervisé . . . . .	39

3.4.3	Apprentissage renforcé . . . . .	40
3.4.4	Apprentissage par rétropropagation d'erreur . . . . .	40
3.4.4.1	Méthode de la descente du gradient . . . . .	40
3.4.4.2	Méthode de Newton . . . . .	43
3.4.4.3	Autres méthodes . . . . .	44

## **Chapitre 4: CONCEPTION DU FILTRE**

### **NEURONAL BASÉ SUR LA STRUCTURE DU FILTRE DE KALMAN..... 46**

4.1	LES RÉSEAUX NEURONAUX ARTIFICIELS DANS LA COM- MANDE . . . . .	47
4.1.1	Réseaux de neurones comme estimateurs . . . . .	47
4.1.1.1	Apprentissage de la dynamique directe du système	47
4.1.1.2	Apprentissage de la dynamique inverse du système	48
4.1.2	Réseau de neurones comme contrôleurs . . . . .	49
4.1.2.1	Commande adaptative directe . . . . .	49
4.1.2.2	Commande adaptative indirecte . . . . .	49
4.1.3	Identification des paramètres par les réseaux neuronaux . . .	51

4.1.4	Autres types de commande par réseaux . . . . .	51
4.2	LE FILTRE NEURONAL . . . . .	52
4.2.1	L'estimateur neuronal . . . . .	52
4.2.2	Calcul de la variation de la sortie estimée par rapport au vecteur des poids ( $\frac{\partial \hat{y}_k}{\partial w_{i,j}}$ ) . . . . .	53
4.2.3	Calcul de la variation de la sortie du réseau par rapport au vecteur des poids ( $\frac{\partial \Gamma_{k-1}}{\partial w_{i,j}}$ ) . . . . .	54
4.2.4	Calcul de la variation des états estimés par rapport au vecteur des poids ( $\frac{\partial \hat{x}_k}{\partial w_{i,j}}$ ) . . . . .	54
4.2.5	Le prédicteur neuronal . . . . .	57
4.2.6	Choix du modèle du réseau . . . . .	58
4.2.7	Choix du nombre de couches du réseau neuronal . . . . .	58

## Chapitre 5:        SIMULATION DU FILTRE

	NEURONAL . . . . .	60
5.1	ESTIMATEUR NEURONAL . . . . .	61
5.1.1	Modèle de l'avion Boeing 747 . . . . .	61
5.1.2	Structure et caractéristiques du réseau de neurones utilisé . . . . .	63

5.1.3	Données de simulation . . . . .	64
5.1.3.1	Les bruits sont blancs et de covariances connues . .	64
5.1.3.2	Les bruits sont non blancs et non gaussiens de co- variances non connues . . . . .	65
5.1.4	Structure du programme d'estimation . . . . .	66
5.2	PRÉDICTEUR NEURONAL . . . . .	68
5.2.1	Le système en étude . . . . .	68
5.2.1.1	Modèle des réservoirs . . . . .	69
5.2.2	Données de Simulation . . . . .	75
5.2.3	Structure et caractéristiques du réseau de neurones utilisé .	81
5.2.4	Structure du programme de prédiction . . . . .	81

## **Chapitre 6: ANALYSE COMPARATIVE DU**

### **FILTRE DE KALMAN ET DE L'ESTIMATEUR**

#### **NEURONAL . . . . . 85**

#### **6.1 ANALYSE COMPARATIVE DES DEUX ESTIMATEURS . . . . . 86**

##### **6.1.1 Les bruits sont blancs et gaussiens de statistiques connues . 86**



6.1.2	Les bruits sont non blancs et non gaussiens de statistiques inconnues . . . . .	93
6.1.2.1	L'intensité des bruits est faible . . . . .	93
6.1.2.2	Le rapport bruit/signal à 5% . . . . .	97
6.1.2.3	Le rapport bruit/signal à 10% . . . . .	102
6.1.3	Effet de la variation des paramètres du système sur les deux estimateurs . . . . .	107
6.1.3.1	Perturbation des pôles de $F$ à 5% dans le cas d'un bruit blanc gaussien . . . . .	111
6.1.3.2	Perturbation à 10% des pôles du système dans le cas d'un bruit blanc gaussien . . . . .	116

## Chapitre 7: ANALYSE COMPARATIVE DU

### FILTRE DE KALMAN ET DU PRÉDICTEUR

#### NEURONAL . . . . . 122

#### 7.1 MESURES COMPLÈTES . . . . . 123

##### 7.1.1 Les bruits sont blancs et gaussiens . . . . . 123

##### 7.1.2 Le rapport bruit/signal à 5% . . . . . 127

##### 7.1.3 Le rapport bruit/signal à 10% . . . . . 132

7.2 MESURES INCOMPLÈTES . . . . .	137
7.2.1 Le rapport bruit/signal à 5% . . . . .	143
7.2.2 Le rapport bruit/signal à 10% . . . . .	148
<b>CONCLUSION GÉNÉRALE . . . . .</b>	<b>154</b>
<b>BIBLIOGRAPHIE . . . . .</b>	<b>158</b>
<b>ANNEXE I . . . . .</b>	<b>167</b>

## Liste des figures

Figure 2.1:	Projection orthogonale (F. L. Lewis, 1986) . . . . .	9
Figure 2.2:	Filtre de Kalman . . . . .	11
Figure 3.1:	Neurone biologique . . . . .	29
Figure 3.2:	Fonctionnement des neurones biologiques . . . . .	30
Figure 3.3:	Propagation de l'influx nerveux dans le neurone biologique	32
Figure 3.4:	Représentation d'un neurone artificiel . . . . .	33
Figure 3.5:	Fonctions d'activation: a) linéaire; b) linéaire avec saturation; c) fonction tout ou rien; d) fonction sigmoïde . . . . .	34
Figure 3.6:	Réseaux multi-couches . . . . .	36
Figure 3.7:	Réseaux connectés . . . . .	37
Figure 3.8:	Apprentissage supervisé . . . . .	39
Figure 3.9:	Apprentissage non supervisé . . . . .	39
Figure 4.1:	Dynamique directe . . . . .	48

Figure 4.2:	Dynamique inverse . . . . .	48
Figure 4.3:	Commande adaptative directe . . . . .	49
Figure 4.4:	Commande adaptative indirecte . . . . .	50
Figure 4.5:	Structure du filtre neuronal . . . . .	53
Figure 4.6:	Organigramme de calcul de $\frac{\partial \hat{x}_1(k)}{\partial w}$ à l'instant $k$ . . . . .	55
Figure 4.7:	Structure du predicteur neuronal . . . . .	57
Figure 5.1:	Structure du Réseau estimateur utilisé . . . . .	64
Figure 5.2:	Signaux d'entrée utilisés pour l'estimation des états du Boeing 747; a) Entrée $u_1$ (rad); b) Entrée $u_2$ (rad). . . . .	66
Figure 5.3:	Organigramme de programmation de l'estimateur neuronal et de comparaison avec le filtre de Kalman . . . . .	67
Figure 5.4:	Réseau prédicteur en phase d'apprentissage. . . . .	68
Figure 5.5:	Réseau prédicteur en rappel. . . . .	69
Figure 5.6:	a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Baskatong) . . . . .	73

Figure 5.7: a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Cabonga) . . . . .	74
Figure 5.8: a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Maniwaki) . . . . .	75
Figure 5.9: Données sur deux ans du réservoir Cabonga utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel . . . .	78
Figure 5.10: Données sur deux ans du réservoir Maniwaki utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel . . . .	79
Figure 5.11: a) Données sur deux ans du réservoir Baskatong utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel . .	80
Figure 5.12: Réseau prédicteur utilisé . . . . .	82
Figure 5.13: Organigramme de programmation du prédicteur neuronal et de comparaison avec le filtre de Kalman . . . . .	83
Figure 6.1: Estimés de l'état $x_1$ (vitesse sur l'axe de roulis) par le filtre neuronal et le filtre de Kalman pour une condition initiale égale à 0.1 ft/sec. . . . .	88
Figure 6.2: a) Etat $x_1$ (vitesse sur l'axe de roulis) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives. . . . .	89

Figure 6.3: a) Etat $x_2$ (vitesse sur l'axe de lacet) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives. . . . .	90
Figure 6.4: a) Etat $x_3$ (vitesse de tangage) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives. . . . .	91
Figure 6.5: a) Etat $x_4$ (angle de tangage) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives. . . . .	92
Figure 6.6: a) Etat $x_1$ (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative. . . . .	94
Figure 6.7: a) Etat $x_2$ (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative . . . . .	95
Figure 6.8: a) Etat $x_3$ ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative. . . . .	96
Figure 6.9: a) Etat $x_4$ (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative. . . . .	97
Figure 6.10: a) Etat $x_1$ (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative. . . . .	99

Figure 6.11: a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative. . . . . 100

Figure 6.12: a) Etat  $x_3$  ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative. . . . . 101

Figure 6.13: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative. . . . . 102

Figure 6.14: a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreur relative. . . . . 104

Figure 6.15: a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreur relative. . . . . 105

Figure 6.16: a) Etat  $x_3$  ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreurs relative. . . . . 106

Figure 6.17: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreurs relative. . . . . 107

Figure 6.18: a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives. . . . . 112

Figure 6.19: a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives. . . . . 113

Figure 6.20: a) Etat  $x_3$  (vitesse de tangage) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives. . . . . 114

Figure 6.21: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives. . . . . 115

Figure 6.22: a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative. 117

Figure 6.23: a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative. 118

Figure 6.24: a) Etat  $x_3$  (vitesse de tangage) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative. . . . 119

Figure 6.25: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative. . . . 120



Figure 7.1: a) Prédiction à observations complètes; b) Erreurs moyennes de prédiction (Baskatong). . . . .	124
Figure 7.2: a) Prédiction à observations complètes; b) Erreurs relatives; rapport bruit/signal à 1% (Maniwaki). . . . .	125
Figure 7.3: a) Prédiction à observations complètes; b) Erreurs relatives; rapport bruit/signal à 1% (Cabonga). . . . .	126
Figure 7.4: a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Baskatong). . . . .	129
Figure 7.5: a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Maniwaki). . . . .	130
Figure 7.6: a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Cabonga). . . . .	131
Figure 7.7: a) Prédiction à observations complètes; rapport bruit/signal à 10%; b) Erreurs relatives (Baskatong). . . . .	134
Figure 7.8: a) Prédiction à observations complètes; rapport signal-bruit à 10%; b) Erreurs relatives.(Maniwaki) . . . . .	135
Figure 7.9: a) Prédiction à observations complètes; rapport signal-bruit à 10%; b) Erreurs relatives (Cabonga). . . . .	136
Figure 7.10: a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Baskatong). . . . .	140

Figure 7.11: a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Maniwaki). . . . .	141
Figure 7.12: a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Cabonga). . . . .	142
Figure 7.13: a) Prédiction à observations incomplètes rapport bruit/signal à 5%; b) Erreurs relatives (Baskatong). . . . .	145
Figure 7.14: a) Prédiction à observations incomplètes rapport bruit/signal à 5%; b) Erreurs relatives (Cabonga). . . . .	146
Figure 7.15: a) Prédiction à observations incomplètes rapport signal-bruit à 5%; b) Erreurs relatives (Maniwaki). . . . .	147
Figure 7.16: a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Baskatong). . . . .	150
Figure 7.17: a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Maniwaki). . . . .	151
Figure 7.18: a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Cabonga). . . . .	152

# Liste des tableaux

Tableau 5.1: Caractéristiques du réseau estimateur utilisé . . . . .	64
Tableau 5.2: Paramètres du modèle des réservoirs (modèle basé sur deux mesures passées) identifiés par le filtre de Kalman étendu . . . . .	71
Tableau 5.3: Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu. . . . .	72
Tableau 5.4: Paramètres d'apprentissage . . . . .	76
Tableau 5.5: Caractéristiques du réseau prédicteur utilisé . . . . .	82
Tableau 7.1: Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu. Mesures bruitées avec un rapport signal-bruit de 5% . . . . .	128
Tableau 7.2: Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu. Mesures bruitées avec un rapport bruit/bruit de 10% . . . . .	133
Tableau 7.3: Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures (les mesures sont traitées) . . . .	139

Tableau 7.4: Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures pour un rapport bruit/signal à 5% 144

Tableau 7.5: Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures pour un rapport bruit/signal à 10% 149

# Liste des symboles

$E[.]$  : Opérateur espérance.

$[.]^t$  : Opérateur transposé.

$\Re$  : Ensemble des nombres réels

$\omega_k$  : Bruit de modèle.

$v_k$  : Bruit de mesure.

$\tilde{y}_{k|k-1}$  : Erreur sur la prédiction (Innovation).

$\Sigma_{k|k}$  : Matrice de covariance a posteriori.

$\Sigma_{k|k-1}$  : Matrice de covariance a priori.

$x_{k|k}$ : Estimé de l'état  $x_k$  à l'instant  $k$  à partir des informations jusqu'à l'instant  $k$

$x_{k|k-1}$ : Estimé de l'état  $x_k$  à l'instant  $k$  à partir des informations jusqu'à l'instant  $k - 1$

$F_k$ : Matrice variante d'état du système

$F$ : Matrice invariante d'état du système

$A$ : Matrice invariante d'état du système

$B_k$ : Matrice variante d'entrée du système

$B$ : Matrice invariante d'entrée du système

$C_k$ : Matrice variante de sortie du système

$C$ : Matrice invariante de sortie du système

$G_k$ : Matrice variante d'entrée de bruit de modèle

$G$ : Matrice invariante d'entrée de bruit de modèle

$Q_k$ : Matrice variante d'auto-covariance du bruit de modèle

$Q$ : Matrice invariante d'auto-covariance du bruit de modèle

$R_k$ : Matrice variante d'auto-covariance du bruit de mesure

$R$ : Matrice invariante d'auto-covariance du bruit de mesure

$K_k$ : Matrice de gain du filtre de Kalman

$S$ : Matrice de covariances des bruits du système

$\delta_{kj}$ : Opérateur de Kronecker

$\xi_k$ : Séquence de bruit blanc.

$\zeta_k$ : Séquence de bruit blanc.

$\nu_k$ : Séquence de bruit blanc.

$\varrho_k$ : Séquence de bruit blanc.

$\theta$ : Vecteur de paramètres.

$w_{ij}$ : Poids pondéré entre le neurone  $j$  et  $i$ .

$b_i$ : Biais du neurone  $i$ .

$w$ : Matrice de poids.

$n_e$ : Nombre d'échantillons (données).

$n_p$ : Longueur du vecteur de sortie.

$n_a$ : Longueur du vecteur d'entrée.

$n_o$ : Nombre de neurones de la couche cachée.

$n_r$ : Nombre de mesures passées.

$J$ : Fonction coût.

$\eta$ : Coefficient d'apprentissage.

$\Gamma$ : Vecteur de sortie du réseau de neurones.

$u_a$ : vitesse sur l'axe de roulis

$w_a$  : vitesse sur l'axe de lacet

$q_a$  : vitesse de tangage

$\theta_a$  : angle de tangage

$\phi$ : matrice entrée/sortie à l'instant  $k$

$\mu$ : moment utilisé dans la loi d'apprentissage de Levenberg Marquardt

$u_1(t)$ : premier signal d'entrée de l'avion Boeing 747

$u_2(t)$ : deuxième signal d'entrée de l'avion Boeing 747

$u_k$ : vecteur des entrées du système

## Liste des annexes

- Annexe I: Démonstration du théorème de l'existence du régime permanent du filtre de Kalman



# Chapitre 1

## INTRODUCTION GÉNÉRALE

Plusieurs questions se sont posées concernant le cerveau de l'homme; par exemple, pourquoi l'humain arrive-t-il à piloter un hélicoptère sans avoir besoin de résoudre des équations complexes, alors que pour simuler le pilotage d'un hélicoptère, il faut solutionner six équations différentielles non linéaires? Pourquoi l'homme est-il capable de reconnaître autant de gens avec qui il vit sans faire un effort particulier? Le désir de répondre à ces questions pousse aujourd'hui des chercheurs à vouloir outrepasser les modèles mathématiques, aux équations particulièrement complexes, pour étudier un autre type de systèmes qui agiraient comme des humains. On assiste désormais à une mutation qui conjugue les efforts de la biologie, des sciences cognitives et des sciences de l'ingénieur.

Dans une première approche les chercheurs portent leurs efforts sur le calcul symbolique basé sur l'hypothèse que le raisonnement, modélisant la pensée, est une combinatoire de symboles soumise à des règles logiques. Très rapidement, il sera démontré que cette approche symbolique a des faiblesses. Une première faiblesse tient à la lenteur des méthodes séquentielles utilisées. La deuxième faiblesse, la plus évidente, concerne l'apprentissage. En dépit des efforts importants de plusieurs équipes de chercheurs, il semble difficile, au niveau symbolique, de simuler l'apprentissage.

Compte tenu de ces limitations de l'approche symbolique, les chercheurs dirigent leurs investigations vers une approche où, d'une part, le traitement des informa-

tions se fait essentiellement en parallèle et la représentation des connaissances est distribuée; et d'autre part, l'apprentissage est facilement réalisé. Ce mode de représentation délocalisé et d'apprentissage possible peut être assuré par les réseaux de neurones artificiels, une approche dite connexioniste.

L'idée des réseaux neuronaux prend de l'ampleur dès les années 1943 lorsque McCulloch (1943) introduisit les neurones simplifiés qui, en fait, étaient des répliques des neurones biologiques en plus d'avoir des composants pouvant exécuter des tâches de calcul. Vers les années 50, Hebb (1954) écrivait le livre *"The Organization of Behavior"* qui ravivait l'idée selon laquelle le conditionnement classique psychologique est omniprésent dans l'animal parce qu'il est une propriété des neurones individuels. De proche en proche, cette idée amena Hebb à trouver une loi d'apprentissage à partir des synapses des neurones. Cette loi a servi de base aux nombreux autres chercheurs pour développer plusieurs théories des réseaux neuronaux.

En 1954, Minsky (1954) construisit le premier neuro-calculateur "Snack". Bien que les résultats produits ne fussent pas parfaitement satisfaisants, ils avaient tout de même permis de développer de nouvelles idées de modélisation qui furent utilisées par d'autres chercheurs. La même année Gabor (1954) inventa "The learning Filter" qui utilisait la descente du gradient pour obtenir les poids optimaux qui minimisent l'erreur quadratique entre le signal de sortie observé et celui généré sur l'information passée.

En 1958, Rosenblatt (1958) inventa le Perceptron qui fut à la base de la découverte du premier neuro-calculateur ayant donné des résultats probants " le Mark I Perceptron".

Vers les années 1959, Widrow, aidé par ses étudiants, découvrit le réseau Adaline (ADaptive LINear Element). Plusieurs applications de l'Adaline furent réalisées.

En 1961, Rosenblatt (1961) proposait la rétropropagation de l'erreur pour l'apprentissage des réseaux multi-couches. Cette approche n'avait pas connu grand succès dans la mesure où elle utilisait les fonctions d'activation non-différentiables. De 1967 à 1982, des recherches progressent et font nettement avancer la théorie et les applications dans le domaine des réseaux de neurones. Hopfield (1982) s'est particulièrement intéressé aux réseaux récurrents dont il a développé les lois d'apprentissage. Kosko (1988) s'est le plus penché sur les réseaux à apprentissage spatio-temporel. Quant à Malsburg (1973), Grossberg (1982), Kohonen (1988) leurs champs de prédilection étaient les réseaux à apprentissage compétitif.

Les réseaux neuronaux artificiels sont une technologie émergente qui présente un champ d'applications extrêmement variées. C'est ainsi que les réseaux sont utilisés dans l'optimisation, la commande des systèmes dynamiques, la modélisation des processus, la compression d'images et des données, le tracé des circuits intégrés, la poursuite de cibles, la reconnaissance des formes, de la parole et des signatures, et de la prédiction.

L'un des outils de prédiction les plus fiables dont disposent les chercheurs actuellement est le filtre de Kalman. Sa fiabilité est liée à l'optimalité de son gain dont le calcul est fonction des statistiques des bruits de modèle et de mesure du système en étude. Les méthodes actuelles de détermination du gain de Kalman supposent que les bruits sont blancs ou blancs et gaussiens et que leurs statistiques sont parfaitement connues. L'utilisation du filtre de Kalman admet par ailleurs que le modèle du système, dont il faut estimer les états, est parfaitement donné. Aussi admettons-nous que la meilleure méthode d'estimation du gain de Kalman, ou de façon plus générale, le meilleur estimateur dans un environnement stochastique, s'il existe, serait celui qui ne tiendrait pas compte des statistiques des bruits mais estimerait

tout de même les états du système. En d'autres termes, nous voulons savoir si on peut trouver une structure qui permettrait de régler le gain quel que soit le type de bruit qui perturbe le système et qui résisterait aux variations des paramètres du système. Les réseaux de neurones artificiels étant de bons estimateurs, ne peuvent-ils pas aider à la conception de ce type de structure ?

Le parallèle que nous sommes amenés à faire entre le filtre de Kalman et les réseaux neuronaux artificiels vient du fait que ces deux structures peuvent faire de l'estimation. Certains auteurs ont mené des études de comparaison entre le filtre de Kalman et les réseaux neuronaux. C'est ainsi que Ruck, Rogers et *al.*(1992) examinent la relation théorique entre les deux approches comparant leur performance sur les données actualisées et s'intéressent au coût de traitement de chaque méthode. Les résultats trouvés montrent que la rétro-propagation a un coût trois fois moindre que le filtre de Kalman en terme de temps de calcul. Singhal et Wu (1989) montrent que l'apprentissage d'un réseau "feed-forward" peut être vu comme un problème d'identification d'un système dynamique non linéaire pouvant se résoudre par l'utilisation de l'algorithme du filtre de Kalman étendu. Ces auteurs font une comparaison entre l'apprentissage par rétro-propagation de l'erreur et l'algorithme de Kalman. De l'analyse des résultats, ils remarquent que l'estimation de Kalman converge plus vite que la rétro-propagation de l'erreur. Toutefois, leur article ne dégage pas une relation entre le filtre de Kalman et la rétro-propagation. Par ailleurs, les données utilisées étaient statiques donc non actualisées. Youji et *al.*(1991) mettent sur pied un nouvel algorithme d'apprentissage basé sur le filtre de Kalman étendu en vue d'estimer les poids du réseau de neurones.

D'autres chercheurs tels que Samir Shah et *al.*(1992), Puskorius, Feldkamp (1994) ont, eux aussi, mis sur pied des lois d'apprentissage des réseaux à propagation avant

en utilisant le filtre de Kalman. Roberts et *al.*, (1997) construisent un estimateur neuronal flou pour estimer les conditions initiales optimales du filtre de Kalman afin de diminuer son temps de convergence.

Actuellement, tous les travaux qui ont trait au filtre de Kalman et aux réseaux neuronaux bâtissent de nouveaux algorithmes d'apprentissage en se servant du filtre de Kalman. Ils ne se sont jamais penchés sur l'estimation de la matrice de gain de Kalman qui, en réalité, constitue la pièce maîtresse du filtre.

Le présent travail cherche à remplacer le gain du filtre de Kalman, dans sa structure, par un réseau de neurones artificiels. Une comparaison des résultats obtenus par cette nouvelle configuration à ceux donnés par le filtre de Kalman dans l'hypothèse d'observations connues et incomplètes montrera que le filtre neuronal est un bon estimateur et un bon prédicteur et qu'il s'adapte le mieux aux variations des paramètres des systèmes. Cependant, une analyse sera faite dans le cas des systèmes très corrompus afin de circonscrire les limites du fonctionnement admissible du filtre.

Pour mener ce travail, nous présentons, au chapitre deux, les informations essentielles sur le filtre de Kalman et quelques méthodes de détermination de son gain. Au troisième chapitre, il s'agit de résumer les notions de base des réseaux neuronaux et quelques unes de leurs méthodes d'apprentissage. Le quatrième chapitre permettra de rappeler brièvement quelques utilisations des réseaux neuronaux en automatique avant de procéder à l'analyse proprement dite de la structure du filtre neuronal qui a pour base le filtre de Kalman. Le gain du filtre de Kalman est effectivement remplacé par un réseau de neurones artificiels (RNA) dont l'ajustement des poids permet de minimiser l'erreur d'estimation. Le chapitre cinq pose les bases de la simulation du filtre neuronal en estimateur et en prédicteur. Dans le cas de

l'estimateur, une application pratique est faite sur le modèle de l'avion Boeing 747 en commande du mouvement longitudinal. En prédiction, le système considéré est composé de trois réservoirs hydrologiques de stockage d'eau de l'Hydro-Québec (Baskatong, Maniwaki et Cabonga). Il s'agit de prédire les mesures des apports naturels aux réservoirs à partir de quelques mesures passées des mêmes apports naturels. Le chapitre six présente les résultats de simulation de l'estimateur neuronal et les compare à ceux donnés par le filtre de Kalman. On observe en outre le comportement des deux estimateurs en présence d'un modèle soumis à des degrés différents de perturbation. Le chapitre sept, quant à lui, présente les résultats de simulation du prédicteur neuronal et du filtre de Kalman en prédisant des mesures des apports naturels des réservoirs Cabonga, Maniwaki et Baskatong. Le cas où les réservoirs sont très corrompus sera aussi considéré afin de voir les limites des performances du prédicteur neuronal. Une conclusion générale résumera les performances du filtre neuronal.

## Chapitre 2

# MÉTHODES DE DÉTERMINATION DU GAIN DU FILTRE DE KALMAN

### Introduction

L'objet principal de ce travail étant d'étudier un filtre neuronal basé sur la structure du filtre de Kalman, il s'avère donc nécessaire de rappeler les grandes étapes de calcul du filtre de Kalman et les propriétés qui déterminent sa stabilité et la convergence des estimés. Nous présenterons enfin quelques méthodes de détermination de son gain.

## 2.1 GÉNÉRALITÉS SUR LE FILTRE DE KALMAN

Soit une réalisation d'un processus aléatoire  $y_k$ ; étant donné un espace de Hilbert  $\ell_2$  muni d'un produit interne et d'une norme.

Considérons le système défini par:

$$x_{k+1} = F_k x_k + B_k u_k + G_k \omega_k \quad (2.1)$$

$$y_k = C_k x_k + v_k \quad (2.2)$$

$x_k \in R^n$  (vecteur d'état du système à l'instant  $k$ ),  $y_k \in R^r$  (vecteur de sortie du système à l'instant  $k$ ),  $\omega_k$  est le bruit de modèle et  $v_k$  le bruit de mesure. Les matrices  $F_k$  (matrice d'état),  $B_k$  (matrice d'entrée),  $G_k$  (matrice d'entrée de bruit de modèle) et  $C_k$  (matrice de sortie) sont de dimensions respectives  $(n \times n)$ ,  $(n \times p)$ ,  $(n \times m)$  et  $(r \times n)$ . Les séquences de bruit  $\omega_k$  et  $v_k$  sont supposées être des séquences de bruit blanc, conjointement gaussiennes de moyennes nulles et de variances respectives  $Q_k$  et  $R_k$ . L'état initial  $x_0$  est supposé être gaussien de moyenne nulle et indépendant des séquences de bruit  $\omega_k$  et  $v_k$ .

L'espace futur  $y_k^+$ , l'espace passé  $y_k^-$  et l'espace global  $y_k$  sont respectivement engendrés par  $\{x_0, \omega_0, \dots, \omega_{k-1}, v_k\}$ ;  $\{x_0, \omega_0, \dots, \omega_{k-1}, v_0, \dots, v_k\}$ ;  $\{x_0, \omega_0, \dots, \omega_{k-1}, v_0, v_1, v_2, \dots\}$ .

Dans cet espace de Hilbert des variables aléatoires, l'opération de corrélation entre les variables aléatoires peut être interprétée comme le produit scalaire.

Il a été démontré par Malhamé (1994) que quelles que soient deux variables aléatoires  $x_1$  et  $x_2$ , les variables aléatoires  $x_1 - E[x_1|x_2]$  et  $x_2$  sont indépendantes; autrement dit:

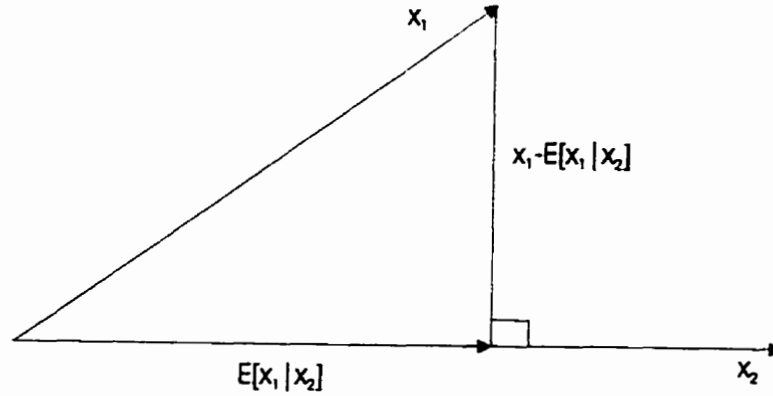
$$E[(x_1 - E(x_1|x_2))|x_2] = 0 \quad (2.3)$$

donc  $x_1 - E(x_1|x_2)$  est orthogonale à  $x_2$ . C'est ce qui est illustré à la figure 2.1.

Le problème du filtrage au sens de Kalman peut être résumé de la manière suivante:

étant donné un système dont l'évolution dynamique est modélisée à l'aide de l'équation (2.1) et dont on observe des combinaisons linéaires de l'état suivant l'équation de mesure (2.2), il faut trouver la meilleure estimation  $\hat{x}$  de l'état  $x$  à l'instant  $k$  à partir des observations effectuées jusqu'à l'instant  $t$  et au sens du critère de la variance conditionnelle minimum.





**Figure 2.1:** Projection orthogonale (F. L. Lewis, 1986)

L'idée fondamentale ici est de décomposer les mesures en segments indépendants. Un premier grand segment comporte toutes les mesures passées  $y_0, y_1, y_2, y_3, \dots, y_{k-1}$  et un deuxième segment est constitué de la dernière mesure  $y_k$ .

On cherche:

$$x_{k|k} = E[x_k | y^k]; \quad (2.4)$$

où  $y^k = [y_0, y_1, y_2, y_3, \dots, y_k]^t$ .

Posons:

$$\tilde{y}_{k|k-1} = y_k - y_{k|k-1} = y_k - E[y_k | y^{k-1}] \quad (2.5)$$

On vérifie que  $y^k$  et  $[(y^{k-1}, \tilde{y}_{k|k-1})]^t$  (où  $\tilde{y}_{k|k-1}$  est l'erreur sur la prédiction (innovation) à une étape de l'observation  $y_k$ ) sont fonction l'une de l'autre et véhiculent la même information.

De plus  $y^{k-1}$  et  $\tilde{y}_{k|k-1}$  sont indépendantes donc orthogonales. Il est à cet effet possible d'écrire:

$$x_{k|k} = E[x_k | y^{k-1}] + E[x_k \tilde{y}_{k|k-1}^t] [\Sigma_{k|k-1}^y]^{-1} \tilde{y}_{k|k-1} \quad (2.6)$$

Le terme  $\Sigma_{k|k-1}^y$ , qui est la covariance de  $\tilde{y}_{k|k-1}$ , nous informe sur la crédibilité de l'information et  $x_k \tilde{y}_{k|k-1}^t$  mesure la corrélation entre  $x_k$  et l'innovation.

Par définition:

$$\Sigma_{k|k} = E[(x_k - x_{k|k})(x_k - x_{k|k})^t] \quad (2.7)$$

cette covariance peut encore se mettre sous la forme:

$$\Sigma_{k|k} = \Sigma_{k|k-1} - E[x_k \tilde{y}_{k|k-1}^t][\Sigma_{k|k-1}^y]^{-1}(E[x_k \tilde{y}_{k|k-1}])^t \quad (2.8)$$

En se basant sur les équations (2.1) (2.2), les expressions (2.6) (2.8) deviennent:

$$x_{k|k} = x_{k|k-1} + \Sigma_{k|k-1} C_k^t [C_k \Sigma_{k|k-1} C_k^t + R_k]^{-1} \tilde{y}_{k|k-1} \quad (2.9)$$

et

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^t [C_k \Sigma_{k|k-1} C_k^t + R_k]^{-1} C_k \Sigma_{k|k-1} \quad (2.10)$$

Si on pose:

$$K_k = \Sigma_{k|k-1} C_k^t [C_k \Sigma_{k|k-1} C_k^t + R_k]^{-1} \quad (2.11)$$

Kalman (1960) montre que la densité conditionnelle de moyenne  $x_{k|k}$  et de variance  $\Sigma_{k|k}$  peut être obtenue à partir des relations ci-dessous:

$$\begin{cases} x_{k+1|k+1} &= x_{k+1|k} + K_{k+1}(y_{k+1} - C_k x_{k+1|k}) \\ \Sigma_{k+1|k} &= F_k \Sigma_{k|k} F_k^t + G_k Q_k G_k^t \\ K_k &= \Sigma_{k|k-1} C_k^t [C_k \Sigma_{k|k-1} C_k^t + R_k]^{-1} \end{cases} \quad (2.12)$$

Le filtre de Kalman, dans sa globalité, peut se résumer de la façon suivante:

- De la connaissance *a priori* de la dynamique du système.

On calcule:

$$\begin{cases} x_{k+1|k} &= F_k x_{k|k} + B_k u_k \\ \Sigma_{k+1|k} &= F_k \Sigma_{k|k} F_k^t + G_k Q_k G_k^t \end{cases} \quad (2.13)$$

On est ici dans le cas prédictif.

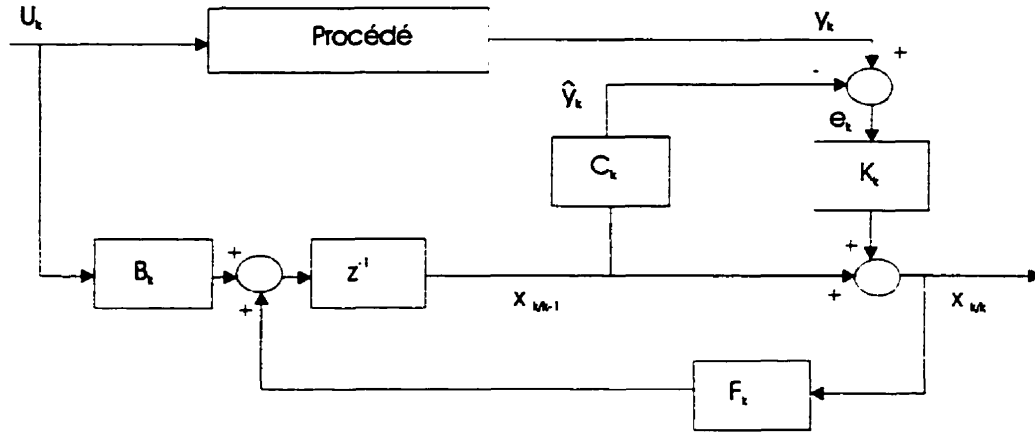


Figure 2.2: Filtre de Kalman

- De la connaissance *a posteriori* (dès que la mesure  $y_{k+1}$  est disponible).

On calcule:

$$\begin{cases} x_{k+1|k+1} = F_k x_{k|k} + B_k u_k + K_{k+1} (y_{k+1} - C_k F_k x_{k|k} + B_k u_k) \\ \Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \Sigma_{k+1|k} C_k^t [C_k \Sigma_{k+1|k} C_k^t + R_k]^{-1} C_k \Sigma_{k+1|k} \\ K_{k+1} = \Sigma_{k+1|k} C_k^t [C_k \Sigma_{k+1|k} C_k^t + R_k]^{-1} \\ \Sigma_{k+1|k+1} = (I - K_k C_k) \Sigma_{k+1|k} \\ x_{k+1|k+1} = x_{k+1|k} + K_{k+1} (y_{k+1} - C_k x_{k+1|k}) \end{cases} \quad (2.14)$$

Le système (2.13) peut encore se mettre sous la forme suivante:

$$\begin{cases} x_{k+1|k} = F_k [I - K_k C_k] x_{k|k-1} + B_k u_k + F_k K_k y_k \\ K_k = \Sigma_{k|k-1} C_k^t [C_k \Sigma_{k|k-1} C_k^t + R_k]^{-1} \\ \Sigma_{k+1|k} = F_k [\Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^t (C_k \Sigma_{k|k-1} C_k^t + R_k)^{-1} C_k \Sigma_{k|k-1}] F_k^t \\ \quad + G_k Q_k G_k^t \end{cases} \quad (2.15)$$

La dernière équation du système ci-dessus est appelée équation de Riccati discrète.

- Initialisation

On pose généralement  $x_{o|o} = \bar{x}_o$  (où  $\bar{x}_o$  est la moyenne de  $x_o$ ).

La matrice de covariance  $\Sigma_{o|o}$  est très souvent égalée à une matrice diagonale définie positive dont les termes diagonaux sont suffisamment grands si on n'est pas certain de la mesure  $y_o$ .

Dans le cas où  $R_k$  est constant, une petite matrice de covariance  $\Sigma_{k|k}$  conduira à un gain de Kalman  $K_k$  faible. D'après la relation (2.9), l'estimation  $x_{k|k}$  dépendra surtout de la précédente soit,  $x_{k|k-1}$ , ce que l'on traduit par une plus grande confiance au modèle par rapport à la mesure. Si  $\Sigma_{k|k}$  est grand, le gain  $K_k$  le sera également et la mise à jour de l'estimation donnera plus d'importance au terme d'innovation; en fait, la mesure qui vient d'arriver. Dans le cas où  $\Sigma_{k|k}$  est constant, pour un  $R_k$  petit, le gain  $K_k$  sera grand, donnant ainsi une plus grande importance à la mesure.

### 2.1.1 Le filtre de Kalman étendu

Pour procéder à l'identification des paramètres des modèles des réservoirs hydrologiques de l'Hydro-Québec au chapitre 6, nous utiliserons le filtre de Kalman étendu. Aussi paraît-il nécessaire de présenter ici l'information essentielle de ce filtre.

Le filtre de Kalman dans sa formulation de base ne peut estimer que les vecteurs d'état des systèmes linéaires. Son application aux systèmes non linéaires nécessite une procédure de linéarisation des systèmes concernés; le filtre ainsi obtenu prend le nom de filtre de Kalman étendu.

Soit un système non linéaire dont les équations d'état et d'observation peuvent se mettre sous la forme:

$$\begin{cases} x_{k+1} &= f_k(x_k) + \omega_k \\ y_k &= h_k(x_k) + v_k \end{cases} \quad (2.16)$$

Il suffit de linéariser comme suit:

$$\begin{cases} F_k &= \left( \frac{\partial f_k}{\partial x} \right)_{x=x_{k|k}} \\ H_k &= \left( \frac{\partial h_k}{\partial x} \right)_{x=x_{k|k-1}} \end{cases}$$

alors le filtre de Kalman étendu est basé sur les équations ci-dessous:

$$\begin{cases} x_{k+1|k} = f_k(x_k) \\ \Sigma_{k+1|k} = F_k \Sigma_{k|k} F_k^t + Q_k \\ K_{k+1} = \Sigma_{k+1|k} H_k^t [H_k \Sigma_{k+1|k} H_k^t + R_k]^{-1} \\ x_{k+1|k+1} = x_{k+1|k} + K_{k+1} (y_{k+1} - H_k x_{k+1|k}) \\ \Sigma_{k+1|k+1} = \Sigma_{k+1|k} - K_{k+1} H_k \Sigma_{k+1|k} \end{cases} \quad (2.17)$$

### 2.1.1.1 Identification des paramètres par le filtre de Kalman

L'une des applications du filtre de Kalman étendu est l'identification adaptative des paramètres des systèmes dynamiques.

Soit un système décrit par les équations ci-dessous:

$$\begin{cases} x_{k+1} = F_k(\theta)x_k + g_k(\theta)\xi_k \\ y_k = C_k(\theta)x_k + \zeta_k \end{cases} \quad (2.18)$$

$x_k \in \mathbb{R}^n; \xi_k \in \mathbb{R}^p; \zeta_k \in \mathbb{R}^q$   $1 \leq p, q \leq n$ .  $\{\xi_k\}$  et  $\{\zeta_k\}$  sont des séquences de bruits blancs gaussiens non corrélés. Les matrices  $F_k(\theta)$ ,  $g_k(\theta)$  et  $C_k(\theta)$  sont fonction du vecteur  $\theta$  inconnu. L'objectif premier est d'identifier  $\theta$ .

Posons:

$$\theta_{k+1} = \theta_k + \varrho_k \quad (2.19)$$

où  $\{\varrho_k\}$  est une séquence de bruit blanc gaussien non corrélé avec  $\zeta_k$  et de variance  $S_k$ .  $\theta$  est ici considéré comme un vecteur aléatoire constant.

On peut reformuler le système (2.18) et l'équation (2.19) afin d'obtenir les équations non linéaires ci-dessous:

$$\begin{bmatrix} x_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} F_k(\theta_k)x_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} g_k(\theta_k)\xi_k \\ \varrho_k \end{bmatrix} \quad (2.20)$$

$$y_k = \begin{bmatrix} C_k(\theta_k) & 0 \end{bmatrix} \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} + \zeta_k \quad (2.21)$$

L'application du formalisme du filtre de Kalman étendu permet d'estimer les états  $x_k$  et le vecteur de paramètres  $\theta_k$  en posant:

$$\begin{bmatrix} x_{k|k} \\ \theta_{k|k} \end{bmatrix} = \begin{bmatrix} x_{k|k-1} \\ \theta_{k|k-1} \end{bmatrix} + K_k (y_k - C_k(\theta_{k|k-1})x_{k|k-1}) \quad (2.22)$$

avec:

$$K_k = P_{k|k-1} \begin{bmatrix} C_k(\theta_{k|k-1}) & 0 \end{bmatrix}^t \left[ \begin{bmatrix} C_k(\theta_{k|k-1}) & 0 \end{bmatrix} P_{k|k-1} \begin{bmatrix} C_k \theta_{k|k-1} & 0 \end{bmatrix}^t + R_k \right]^{-1} \quad (2.23)$$

$$\begin{aligned} P_{k|k-1} = & \begin{bmatrix} F_{k-1}(\theta_{k-1|k-1}) & \frac{\partial}{\partial \theta} [F_{k-1}(\theta_{k-1|k-1})x_{k-1|k-1}] \\ 0 & I \end{bmatrix} P_{k-1|k-1} \\ & \cdot \begin{bmatrix} F_{k-1}(\theta_{k-1|k-1}) & \frac{\partial}{\partial \theta} [F_{k-1}(\theta_{k-1|k-1})x_{k-1|k-1}] \\ 0 & I \end{bmatrix}^t \\ & + \begin{bmatrix} g_{k-1}(\theta_{k-1|k-1})Q_{k-1}g_{k-1}^t(\theta_{k-1|k-1}) & 0 \\ 0 & S_{k-1} \end{bmatrix} \end{aligned} \quad (2.24)$$

et

$$P_{k|k} = \begin{bmatrix} I - g_k[(\theta_{k|k-1}) & 0] \end{bmatrix} P_{k|k-1} \quad (2.25)$$

### 2.1.2 Stabilité et Convergence du Filtre de Kalman

Nous supposons que le modèle considéré est invariant dans le temps et est défini par le système ci-dessous:

$$\begin{cases} x_{k+1} = Fx_k + Bu_k + G\omega_k \\ y_k = Cx_k + v_k \end{cases} \quad (2.26)$$

$x_k \in R^n$ ,  $y_k \in R^r$ ,  $\omega_k$  est le bruit de modèle et  $v_k$  le bruit de mesure. Les matrices  $F$ ,  $B$ ,  $G$  et  $C$  sont de dimensions respectives  $(n \times n)$ ,  $(n \times p)$ ,  $(n \times m)$  et  $(r \times n)$ . Les séquences de bruit  $\omega_k$  et  $v_k$  sont supposées être des séquences de bruit blanc, conjointement gaussiennes de moyennes nulles et de variances respectives  $Q$

et  $R$ . L'état initial  $x_0$  est supposé être gaussien de moyenne nulle et indépendant des séquences de bruit  $\omega_k$  et  $v_k$ .

L'un des problèmes dont il faut tenir compte lors de la conception du filtre de Kalman est la garantie de la stabilité asymptotique de son erreur d'estimation; autrement dit, s'assurer de la convergence de l'estimé  $x_{k|k}$  vers la valeur réelle  $x_k$ . La solution à ce problème revient à étudier la dynamique de l'erreur et à trouver ses conditions de stabilité.

Soit l'erreur d'estimation a priori suivante:

$$\begin{cases} \tilde{x}_{k+1|k} &= x_{k+1} - x_{k+1|k} \\ &= F(I - K_k C)\tilde{x}_{k|k-1} - FK_k v_k + G\omega_k \end{cases} \quad (2.27)$$

La stabilité asymptotique de  $\tilde{x}_{k+1|k}$  dépend de la matrice  $F(I - K_k C)$ . Par ailleurs, rappelons que la paire  $(F, C)$  est observable si les pôles de  $(F - LC)$  peuvent être placés arbitrairement par un choix approprié de la matrice  $L$ .  $(F, C)$  est détectable si  $(F - LC)$  peut être asymptotiquement stable pour un choix de  $L$ .  $(F, B)$  est atteignable si les pôles de  $(F - BK)$  peuvent être placés arbitrairement par un choix approprié de la matrice  $K$ ; de même  $(F, B)$  est stabilisable si le choix d'une matrice  $K$  peut rendre  $(F - BK)$  asymptotiquement stable.

Le théorème suivant (P. R. Kumar et P. Varaiya, 1986) permet de dégager les conditions de stabilité et de convergence du filtre de Kalman.

#### **Théorème (existence du régime permanent)**

*Soit  $Q = \sqrt{Q}\sqrt{Q}^t \geq 0$  la covariance du bruit de modèle du procédé et  $R > 0$  la covariance du bruit de mesure. Supposons  $(F, G\sqrt{Q})$  atteignable;  $(F, C)$  est détectable si et seulement si:*

*a) l'équation de Riccati a une solution unique  $\bar{\Sigma}$  indépendante de  $\bar{\Sigma}_{0|-1}$ ;*

b) l'erreur d'estimation a priori, avec pour gain  $\bar{K} = F\bar{L}$  où

$$\bar{L} = \bar{\Sigma}C^t[C\bar{\Sigma}C^t + R]^{-1} \quad (2.28)$$

est asymptotiquement stable et

$$\lim_{k \rightarrow \infty} \Sigma_{k+1/k} = \bar{\Sigma} \quad (2.29)$$

La démonstration de ce théorème se trouve à l'annexe A.

Le théorème spécifie que si l'état du système est atteignable par le bruit du système (le bruit excite suffisamment les modes du système), le filtre de Kalman est asymptotiquement stable si  $(F, C)$  est détectable. Ainsi pour garantir la stabilité du filtre, il faut bien choisir la matrice de sortie  $C$  et s'assurer que le système est suffisamment bruité.

Le filtre de Kalman (figure 2.2) est reconnu comme étant le meilleur estimateur linéaire dans un environnement stochastique lorsque les variances des bruits sont connues. Si de plus les bruits blancs sont gaussiens, il est optimal. Il pose néanmoins un problème fondamental dans la mesure où ses performances sont liées à la parfaite connaissance des statistiques des bruits de mesure, des statistiques des bruits de modèle et des paramètres du système. Ces statistiques, dans la plupart des cas, sont mal connues, ce qui conduit plusieurs chercheurs à diriger leurs investigations sur la détermination du gain du filtre de Kalman à partir du vecteur d'informations entrée/sortie du système.



## 2.2 QUELQUES MÉTHODES DE DÉTERMINATION DU GAIN DU FILTRE DE KALMAN

### 2.2.1 Méthodes indirectes

L'essentiel de ces méthodes s'orchestre autour de l'identification des covariances des bruits desquelles est déduit le gain  $K$ . Plusieurs formalismes sont présentés dans la littérature et nous en expliquons ici quelques-uns.

#### 2.2.1.1 Méthode de corrélation de Koji Tajima

Cette méthode permet d'estimer le gain du filtre de Kalman à partir des matrices d'auto-covariance  $Q$ ,  $R$  et de la matrice de covariance  $S$ .

Soit le système linéaire stochastique suivant:

$$\begin{cases} x_{k+1} = Fx_k + G\omega_k \\ y_k = Cx_k + \nu_k \end{cases} \quad (2.30)$$

$x_k \in R^n$ ,  $y_k \in R^r$ ,  $\omega_k$  et  $\nu_k$  sont des bruits blancs gaussiens de variances:

$$\begin{cases} E[\omega_k \omega_j^t] = Q\delta_{kj} \\ E[\nu_k \nu_j^t] = R\delta_{kj} \\ E[\omega_k \nu_j^t] = S\delta_{kj} \end{cases} \quad (2.31)$$

Koji Tajima (1978) suppose que le système est observable et se propose d'identifier  $Q$ ,  $R$ , et  $S$  à partir des observations  $y_k$ .

Le modèle ARMA du système ci-dessus s'écrit:

$$\sum_{i=0}^m a_i y_{k-i} = \sum_{i=0}^m H A_{i+1} \omega_{k-i-1} + \sum_{i=0}^m a_i \nu_{k-i} \quad (2.32)$$

avec:

$$\begin{cases} A_i = FA_{i-1} + a_{i-1}G & i = 2, 3, \dots, m \\ A_1 = a_0G \end{cases}$$

et les  $a_i$  sont les coefficients du polynôme caractéristique de  $F$ . En posant:

$$Z_k = \sum_{i=0}^m a_i y_{k-i} \quad (2.33)$$

on a:

$$Z_k = \sum_{i=0}^m C A_{i+1} \omega_{k-i-1} + \sum_{i=0}^m a_i \nu_{k-i} \quad (2.34)$$

Ce qu'il faut remarquer, c'est que les autocovariances de  $Z_k$  sont toutes nulles pour  $i > m$ ; celles qui sont non nulles sont fonction de Q, R et S seulement. Soit:

$$L_i = E[Z_k Z_{k-i}^t] \quad (2.35)$$

la matrice de covariance de  $Z_k$  au retard  $i$  et d'après l'expression de  $Z_k$ , on peut écrire:

$$\left\{ \begin{array}{l} L_0 = E[Z_k Z_k^t] \\ \quad = C(A_1 Q A_1^t + A_2 Q A_2^t + \dots + A_m Q A_m^t) C^t \\ \quad \quad + (a_0^2 + a_1^2 + \dots + a_m^2) R \\ \quad \quad + C(a_1 A_1 + a_2 A_2 + \dots + a_m A_m) S \\ \quad \quad + S^t(a_1 A_1^t + a_2 A_2^t + \dots + a_m A_m^t) C^t \\ \quad \cdot \\ \quad \cdot \\ L_i = C(A_{i+1} Q A_1^t + A_{i+2} Q A_2^t + \dots + A_{i+m} Q A_m^t) C^t \\ \quad \quad + (a_i a_0 + a_{i+1} a_1 + \dots + a_m a_{m-i}) R \\ \quad \quad + C(a_0 A_i + a_1 A_{i+1} + \dots + a_{m-i} A_m) S \\ \quad \quad + S^t(a_{i+1} A_1^t + a_{i+2} A_2^t + \dots + a_m A_{m-i}^t) C^t \\ \quad \cdot \\ \quad \cdot \\ L_m = a_m R + a_0 C A_m S \end{array} \right. \quad (2.36)$$

On pose:

$$\hat{L}_i = \frac{1}{N} \sum_{k=1}^N Z_k Z_{k-i}^t \quad i = 0, \dots, m \quad (2.37)$$

où  $\hat{L}_i$  est l'estimé de  $L_i$  et N le nombre d'échantillons utilisés. Pour obtenir les covariances Q, R, et S il suffit de résoudre le système d'équations (2.36) où les  $L_i$

sont remplacés par leurs estimés  $\hat{L}_i$ . Généralement, on s'arrange de manière à ce que le nombre d'équations du système (2.36) soit égal au nombre d'inconnues. Si tel n'est pas le cas, on pose des restrictions sur la forme des matrices Q, R et S pour diminuer le nombre d'inconnues. Ayant déterminé Q, R et S, la conception du filtre de Kalman se fait aisément.

### 2.2.1.2 Méthode de Bergman et Delleur

La méthode de Bergman et Delleur (1985) est basée sur la linéarisation de l'équation non linéaire de mesure autour du dernier estimé du vecteur inconnu  $\alpha$  et sur l'application des conditions d'optimalité du filtre de Kalman. L'on tiendra compte, entre autres, que la séquence d'innovation  $\nu_k$  est un bruit blanc.

L'équation d'observation s'écrit:

$$y_k = f_k(\alpha) + \nu_k \quad (2.38)$$

où  $f_k$  est une fonction non linéaire de  $\alpha$ . Si  $\hat{\alpha}_k$  est l'estimé proche de  $\alpha$ ,  $f_k(\alpha)$  peut être approximée par une série de Taylor d'ordre "un" et l'équation (2.38) s'écrit:

$$y_k \approx f_k(\hat{\alpha}_k) + \frac{\partial}{\partial \alpha} f_k(\hat{\alpha}_k) [\alpha_k - \hat{\alpha}_k] + \nu_k \quad (2.39)$$

ce qui permet de déduire l'équation récurrente ci-dessous:

$$\hat{\alpha}_{k+1} = \hat{\alpha}_k + O_k [y_k - f_k(\hat{\alpha}_k)] \quad (2.40)$$

où  $O_k$  est une séquence de gain. On pose généralement:

$$O_k = S_{k-1} m_k [m_k^t S_{k-1} m_k + 1]^{-1} \quad (2.41)$$

avec:

$$\begin{cases} m_k^t &= \frac{\partial}{\partial \alpha} f_k(\hat{\alpha}_k) \\ &= (\frac{\partial}{\partial \alpha_1} f_k, \frac{\partial}{\partial \alpha_2} f_k, \frac{\partial}{\partial \alpha_3} f_k, \dots, \frac{\partial}{\partial \alpha_s} f_k) / \hat{\alpha}_k \\ S_k &= (I - O_k m_k^t) S_{k-1} \end{cases} \quad (2.42)$$

En posant  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_s)^t$  les éléments inconnus des matrices de covariance  $Q$  et  $R$  et définissant:

$$f_k^i(\alpha) = E(\nu_k \nu_{k-i}) \quad (2.43)$$

on peut obtenir une équation de mesure similaire à celle de l'expression (2.38) et donnée par:

$$\nu_k \nu_{k-i} = f_k^i(\alpha) + v_k \quad (2.44)$$

$i$  est le retard entre les innovations  $\nu_k$ . A chaque instant  $k$ , on a  $p+1$  observations  $\nu_k \nu_{k-i}$ . On peut donc voir que:

$$\hat{\alpha}_k(i) = \hat{\alpha}_k(i-1) + O_k(i)[\nu_k \nu_{k-i} - f_k^i(\hat{\alpha}_k(i-1))] \quad (2.45)$$

Les conditions d'optimalité du filtre de Kalman nous permettent d'écrire:

$$f_k^i(\hat{\alpha}_k(i)) = \begin{cases} C_k \sum_{k|k-1} C_k^t + R_k, & i = 0 \\ 0 & i > 0 \end{cases} \quad (2.46)$$

Pour  $i = 0, 1$

$$\begin{cases} \hat{\alpha}_k(0) = \hat{\alpha}_{k-1}(1) + O_k(0)[\nu_k \nu_k - C_k \sum_{k|k-1} C_k^t + R_k] \\ \hat{\alpha}_k(1) = \hat{\alpha}_k(0) + O_k(1)[\nu_k \nu_{k-1}] \end{cases} \quad (2.47)$$

L'équation récurrente du gain est donnée par:

$$\begin{cases} O_k(i) = S_k(i-1)m_k(i)[m_k^t(i)S_k(i-1)m_k(i) + 1]^{-1} \\ m_k^t(i) = \frac{\partial f_k^i(\alpha_k)}{\partial \alpha} = \frac{\partial E[\nu_k \nu_{k-i}]}{\partial \alpha} \\ S_k(i) = [I - O_k(i)m_k^t(i)]S_k(i-1) \end{cases} \quad (2.48)$$

Rappelons que:

$$\begin{cases} \frac{\partial f_k^i(\alpha)}{\partial \alpha} = C_k \frac{\partial \sum_{k|k-1} C_k^t}{\partial \alpha} + \frac{\partial R}{\partial \alpha} & i = 0 \\ = C_k A_{k-1} [\prod_{l=k-i+1}^{k-1} (I - K_l H_l) A_{k-1}] [(I - K_{k-1} H_k) \frac{\partial \sum_{k-i|k-1} C_k^t}{\partial \alpha} - K_{i-1} \frac{\partial R}{\partial \alpha}] & i > 0 \end{cases} \quad (2.49)$$

avec:

$$\begin{cases} \frac{\partial}{\partial \alpha_j} \sum_{k|k} = (I - K_{k-1} C_k) \frac{\partial}{\partial \alpha_j} \sum_{k|k-1} \\ \frac{\partial}{\partial \alpha_j} \sum_{k+1|k} = F_{k-1} \frac{\partial}{\partial \alpha_j} \sum_{k|k} F_{k-1}^t + g \frac{\partial}{\partial \alpha_j} Q g^t \end{cases} \quad (2.50)$$

### 2.2.1.3 Méthode de Assaf et Quick

Assaf et Quick (1991) proposent de considérer la covariance de l'erreur  $R_k$  comme étant égale au carré d'un certain pourcentage de la mesure. Ils posent en effet:

$$R_k = (0.15y_k)^2 \quad (2.51)$$

Cette relation donne la plus petite somme des carrés des résidus. D'une manière similaire, ils calculent la matrice de covariance du bruit de modèle de la manière suivante:

$$\Sigma_{k|k-1}(i, j) = (0.1A_{ij})^2 \delta_{ij} \quad (2.52)$$

Cette méthode, bien qu'empirique, donne néanmoins des résultats convergents.

### 2.2.1.4 Méthode de Todini

Todini (1978) suppose que les moyennes des bruits de modèle et de mesure ne sont pas nulles et estime leurs statistiques comme ci-dessous:

$$\begin{cases} \bar{V}_k = (\frac{k-1}{k})\bar{V}_{k-1} + \frac{1}{k}(y_k - C_k x_{k|k-1}) \\ \bar{W}_k = (\frac{k-1}{k})\bar{W}_{k-1} + \frac{1}{k}K \nu_k \\ R_k = (\frac{k-1}{k})R_{k-1} + \frac{1}{k}(\nu_k \nu_k^t - C_k \sum_{k|k-1} C_k^t \\ Q_k = (\frac{k-1}{k})Q_{k-1} + \frac{1}{k}(K \nu_k \nu_k^t K^t + \sum_{k|k} - \sum_{k|k-1}) \end{cases} \quad (2.53)$$

Il y a lieu de noter ici que la convergence de ces variables n'est pas du tout garantie, ce qui raréfie l'utilisation de ces équations récurrentes pour l'estimation de ces statistiques.

### 2.2.1.5 Méthode de Mayers et Tapkey

Mayers et Tapley (1976) ont leur méthode basée sur celle de Todini. La différence majeure est due au fait qu'au lieu d'effectuer une moyenne sur toutes les mesures à chaque pas de temps  $k$  jusqu'à l'instant  $t$ , on considère une fenêtre mobile d'une

certaines grandeurs  $l_r$  (pour les statistiques de bruit de mesure) et  $l_q$  (pour celles de bruit de modèle). Les premières estimations des bruits de mesure et de modèle sont affectées d'un poids  $r_j$  et  $q_j$  par le souci de réduire leur influence dans les premiers pas d'échantillonnage.

## 2.2.2 Méthodes directes

Ces méthodes identifient le gain  $K$  à partir du vecteur d'informations entrée/sortie du système. Dans ce paragraphe, nous présenterons seulement la méthode de Jer-Nan Juang et *al.* (1993).

Soit le système linéaire discret stochastique ci-dessous:

$$\begin{cases} x_{k+1} = Fx_k + Bu_k + \omega 1_k \\ y_k = Cx_k + Du_k + \omega 2_k \end{cases} \quad (2.54)$$

où  $\omega 1_k$  et  $\omega 2_k$  sont respectivement les bruits de modèle et de mesure, indépendants, gaussiens et de moyenne nulle. L'estimateur linéaire à variance minimale s'écrit:

$$\begin{cases} x_{k+1|k+1} = Fx_{k|k} + Bu_k + K\epsilon_k \\ y_{k|k} = Cx_{k|k} + Du_k \end{cases} \quad (2.55)$$

avec

$$\begin{cases} \epsilon_k = y_k - y_{k|k} \\ u = [u_0, u_1, \dots, u_N] \\ y = [y_0, y_1, \dots, y_N] \end{cases} \quad (2.56)$$

Jer-Nan Juang et *al.* (1993) ont mené l'étude que nous présentons ici.

Le filtre de Kalman étant linéaire, on peut admettre que son équation est la superposition d'un système déterministe excité par  $u_k$  et d'un système stochastique excité par le bruit. L'équation (2.55) peut donc se décomposer et l'on obtient les deux systèmes ci-dessous (le premier étant déterministe le second stochastique):

$$\begin{cases} x1_{k+1|k+1} = Fx1_{k|k} + Bu_k \\ y1_{k|k} = Cx1_{k|k} + Du_k \end{cases} \quad (2.57)$$

$$\begin{cases} x_{2_{k+1}|k+1} = Ax_{2_{k|k}} + K\epsilon_k \\ y_{2_{k|k}} = Cx_{2_{k|k}} + \epsilon_k \end{cases} \quad (2.58)$$

Si on pose  $x_{0|0} = 0$ , (par simplification posons  $y_{1_k} = y_{1_{k|k}}$ ;  $y_{2_k} = y_{2_{k|k}}$ ) on détermine la sortie  $y_{1_k}$  de la manière suivante:

$$\begin{cases} x_{1_{1|1}} = Bu_0 \\ x_{1_{2|2}} = FBu_0 + Bu_1 \\ x_{1_{3|3}} = F^2Bu_0 + FBu_1 + Bu_2 \\ \vdots \\ \vdots \\ \vdots \\ x_{1_{k|k}} = \sum_{i=1}^k F^{i-1}Bu_{k-i} \end{cases} \quad (2.59)$$

Sachant que:

$$y_{1_{k|k}} = Cx_{1_{k|k}} + Du_k \quad (2.60)$$

on déduit la valeur de la sortie déterministe  $y_{1_k}$  en fonction des matrices F,B,C et D

$$y_{1_{k|k}} = \sum_{i=1}^k CF^{i-1}Bu_{k-i} + Du_k \quad (2.61)$$

Les paramètres  $CF^{i-1}B$  sont appelés paramètres de Markov.

La composante stochastique de la sortie se calcule en posant:

$$y_{2_k} = y_k - y_{1_k} \quad (2.62)$$

Cette composante est appelée résidu de sortie.

La simplification apportée à ce problème permet de calculer K à partir de l'équation (2.58) et connaissant F, C et  $y_{2_k}$ . On a par ailleurs:

$$\epsilon_k = y_{2_k} - Cx_{2_{k|k}} \quad (2.63)$$

Cette nouvelle expression de  $\epsilon_k$  provient du fait qu'en l'absence de bruit (cas déterministe), le filtre ne calcule pas d'innovation sur la valeur de l'estimé;

autrement dit, à chaque instant, l'estimé est le même et les covariances sont nulles; l'erreur qu'on observe donc dans le système provient tout naturellement de la partie stochastique du système, d'où l'expression de l'erreur donnée ci-dessus.

En remplaçant  $\epsilon_k$  par sa valeur dans l'expression (2.58), on obtient le système ci-dessous:

$$\begin{cases} x_{2k+1|k+1} = (F - KC)x_{2k|k} + Ky_{2k} \\ y_{2k} = Cx_{2k|k} + \epsilon_k \end{cases} \quad (2.64)$$

Par la même procédure utilisée à l'équation (2.59) on peut constater que

$$x_{2k|k} = \sum_{i=1}^k (F - KC)^{i-1} Ky_{2k-i} \quad (2.65)$$

ce qui nous permet d'écrire:

$$y_{2k} = \sum_{i=1}^k C(F - KC)^{i-1} Ky_{2k-i} + \epsilon_k \quad (2.66)$$

Posons:

$$\alpha_k = C(F - KC)^{k-i} K \quad k = 1, 2, \dots, p \quad (2.67)$$

La matrice  $F - KC$  est appelée matrice du filtre de Kalman et doit être asymptotiquement stable.

Pour une valeur assez élevée de  $p$ ,  $(F - KC)^{i-1}$  tend vers zéro; autrement dit, on aura:

$$\begin{cases} \alpha_0 \neq 0 \\ \alpha_1 \neq 0 \\ \vdots \\ \alpha_p \neq 0 \\ \alpha_{p+1} = 0 \end{cases} \quad (2.68)$$

La nouvelle expression de  $y_{2k}$  devient:

$$y_{2k} = \sum_{i=1}^p \alpha_i y_{2k-i} + \epsilon_k \quad (2.69)$$



Sous forme matricielle, l'équation (2.69) se transforme comme ci-dessous:

$$y2 = \alpha Y + \Upsilon \quad (2.70)$$

$$\Upsilon = [\epsilon_0 \ \epsilon_1 \ \dots \ \epsilon_N] \quad (2.71)$$

$$\begin{cases} y2 = [y2_0 \ y2_1 \dots y2_p \dots y2_N] \\ \alpha = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_p] \end{cases} \quad (2.72)$$

$$Y = \begin{bmatrix} y2_0 & y2_1 & \cdot & \cdot & \cdot & y2_p & y2_{p+1} & \cdot & \cdot & \cdot & y2_N \\ & y2_0 & \cdot & \cdot & \cdot & y2_{p-1} & y2_{p-2} & \cdot & \cdot & \cdot & y2_{N-1} \\ & & \cdot & & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & & & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ & & & & & y2_0 & y2_1 & \cdot & \cdot & \cdot & y2_{N-p} \end{bmatrix} \quad (2.73)$$

A partir de la relation:

$$y2 = \alpha Y + \Upsilon \quad (2.74)$$

on peut voir  $\alpha$  comme la matrice des paramètres à identifier. La méthode des moindres carrés permet de déterminer la matrice  $\alpha$ . Il faut chercher les valeurs de  $\alpha$  qui minimisent le carré de l'erreur de sortie.

Définissons:

$$J(\alpha) = \sum_{k=0}^N \epsilon^2(k, \alpha) \quad (2.75)$$

soit

$$J(\alpha) = \epsilon^t \epsilon \quad (2.76)$$

on cherche à déterminer  $\hat{\alpha}$  tel que:

$$J(\hat{\alpha}) \leq J(\alpha) \quad (2.77)$$

La méthode des moindres carrés donne:

$$\hat{\alpha} = (Y^T Y)^{-1} Y^t y2 \quad (2.78)$$

Connaissant la valeur de  $\alpha$  (identifié par les moindres carrés), on détermine le gain  $K$  de la manière suivante:

Soit  $\beta_k$  les paramètres de Markov du gain du filtre de Kalman

$$\beta_k = CF^{k-1}K \quad k = 1, 2, \dots, p \quad (2.79)$$

on remarque que:

$$\begin{cases} \beta_1 = CK = \alpha_1 \\ \beta_2 = CFK = \alpha_2 + \sum_{i=1}^1 \beta_{2-i} \alpha_i \\ \beta_3 = CF^2K = \alpha_3 + \sum_{i=1}^2 \beta_{3-i} \alpha_i \\ \vdots \\ \beta_k = \alpha_k + \sum_{i=1}^{k-1} \beta_{k-i} \alpha_i \end{cases} \quad (2.80)$$

Connaissant les valeurs de  $F$ ,  $C$  et ayant déjà identifié  $\alpha_k$ , on peut calculer  $\beta_k$  par les relations données des équations du système (2.80). La valeur du gain  $K$  se déduit par:

$$K = (V^t V)^{-1} V^t Y_k \quad (2.81)$$

avec:

$$V = \begin{bmatrix} C \\ CF \\ \vdots \\ CF^{p-1} \end{bmatrix}; \quad Y_k = \begin{bmatrix} CK \\ CFK \\ \vdots \\ CF^{p-1}K \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad (2.82)$$

## Conclusion

Toutes les méthodes de détermination du gain du filtre de Kalman repertoriées dans ce chapitre supposent que les bruits de modèle et de mesure sont blancs. Cette hypothèse est très limitative dans la mesure où généralement les bruits sont non blancs en réalité. Dans la littérature, pour estimer les statistiques des bruits non

blancs, on procède d'abord à un processus de blanchiment qui, très souvent, fait perdre au bruit presque toutes ses caractéristiques réelles. Aussi proposons-nous un nouveau filtre à base de réseaux de neurones artificiels. Ce filtre, comme celui de Kalman, peut être un estimateur ou un prédicteur. Sa structure et les équations qui la caractérisent sont développées au chapitre 4.

## Chapitre 3

# NOTIONS DE BASE SUR LES RÉSEAUX NEURONAUX

### Introduction

Ce chapitre permet d'avoir une idée du fonctionnement et des propriétés communes des réseaux de neurones aussi bien biologiques qu'artificiels. Il met l'accent sur les différents types d'apprentissage des RNA. C'est en connaissant le fonctionnement des RNA qu'il sera possible de choisir la structure du filtre neuronal et le type du réseau le constituant.

### 3.1 FONDEMENTS BIOLOGIQUES

Le fonctionnement du neurone artificiel est semblable à celui d'un neurone biologique. La maîtrise de l'analyse des réseaux neuronaux biologiques est un atout important pour la compréhension des réseaux neuronaux artificiels (RNA).

Le neurone est l'élément de base du cerveau humain. Il se présente comme illustré à la figure 3.1 et se compose essentiellement du corps cellulaire (ou soma), des dendrites, des synapses et de l'axone.

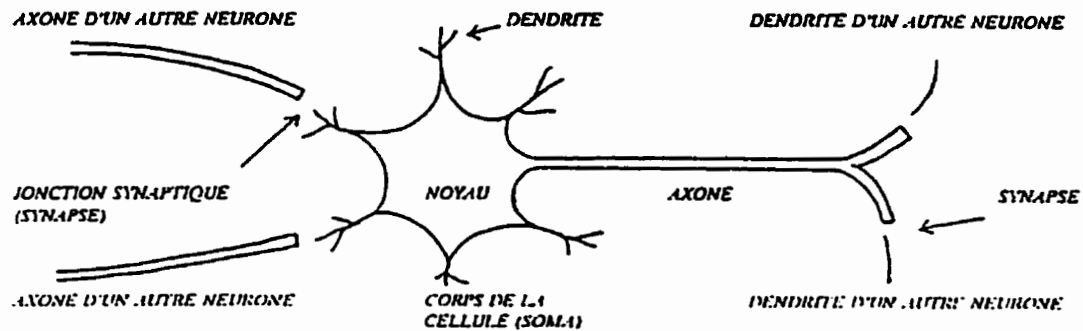


Figure 3.1: Neurone biologique

### 3.1.1 Dendrites

Les dendrites sont des fines extensions tubulaires qui sont essentiellement les récepteurs principaux du neurone pour capter les signaux qui lui parviennent.

### 3.1.2 Le corps cellulaire

Le corps cellulaire effectue les transformations biochimiques nécessaires à la synthèse des enzymes et autres molécules qui assurent la vie du neurone. Il contient le noyau. La forme du corps cellulaire est généralement pyramidale ou sphérique dépendant de sa position dans le cerveau (Naïm, 1990).

### 3.1.3 Synapses

Ce sont les endroits où se font les connexions entre deux neurones. Elles constituent le lieu de transmission de l'information entre les neuro-transmetteurs et les neuro-récepteurs.

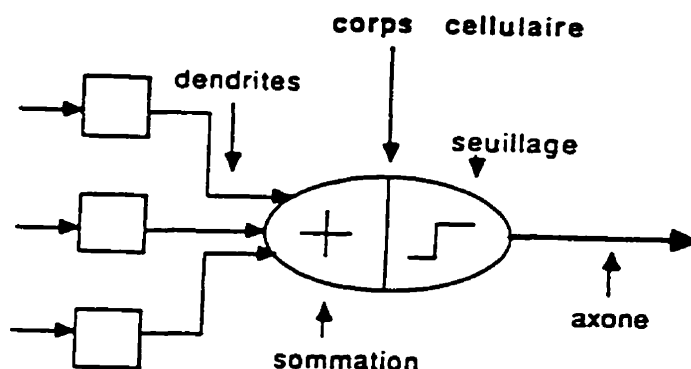


Figure 3.2: Fonctionnement des neurones biologiques

### 3.1.4 L'axone

L'axone sert de moyen de transport des signaux (de l'information) entre neurones. Il se distingue des dendrites par sa forme et par les propriétés de sa membrane externe. En somme, il est généralement plus long que les dendrites (sa longueur varie d'un millimètre à plus d'un mètre). L'axone se ramifie à son extrémité là où il communique avec d'autres neurones alors que les ramifications des dendrites se produisent plutôt près du corps cellulaire.

### 3.1.5 Fonctionnement des neurones biologiques

Le corps cellulaire se comporte comme un sommateur à seuil (figure 3.2). Si la somme des informations reçues des dendrites est supérieure à un seuil donné, le neurone s'excite et l'information résultante est transmise à d'autres neurones par l'axone; c'est la genèse de l'influx nerveux. Si l'information ne dépasse pas le seuil, le neurone ne s'excite pas et reste inactif. Il faut rappeler ici que l'information est

sous forme de potentiel électrique et le seuil est généralement de  $-70mV$ .

La propagation de l'influx nerveux, généré par le corps cellulaire du neurone, est assurée par un mécanisme de pompes et de canaux présents dans la membrane axonique. Les pompes considérées sont des pompes à ions qui maintiennent dans l'axone une différence de concentration en ions sodium( $N_a^+$ ) et potassium( $K^+$ ) par rapport au milieu extérieur. Leur rôle est d'expulser les ions sodium et d'attirer les ions potassium.

Au repos, quand aucun influx nerveux n'est transmis, les canaux sont fermés et les pompes assurent le maintien du potentiel négatif dans la cellule et l'axone. Quand il y a propagation d'un influx nerveux, la différence de potentiel entre le milieu extérieur et l'axone diminue. Cela provoque l'ouverture des canaux à sodium situés juste en avant de l'influx nerveux (figure 3.3). Le résultat de l'entrée des ions sodium dans l'axone est une inversion du potentiel qui existait et donc une propagation de l'influx nerveux. Juste après l'ouverture des canaux à sodium, ce sont les canaux à potassium qui s'ouvrent automatiquement. Ce phénomène provoque une ré-inversion du potentiel et ramène l'axone à son état électrique initial. C'est ainsi que se propage de lui-même le potentiel d'action. Dès que le potentiel d'action arrive au niveau des synapses, il y a libération d'une substance chimique appelée neuro-transmetteur qui se fixe sur les récepteurs spécifiques: les neuro-récepteurs situés sur la terminaison du neurone cible. Cette fixation provoque l'ouverture de canaux ioniques, ce qui peut donner naissance à un nouveau signal électrique.

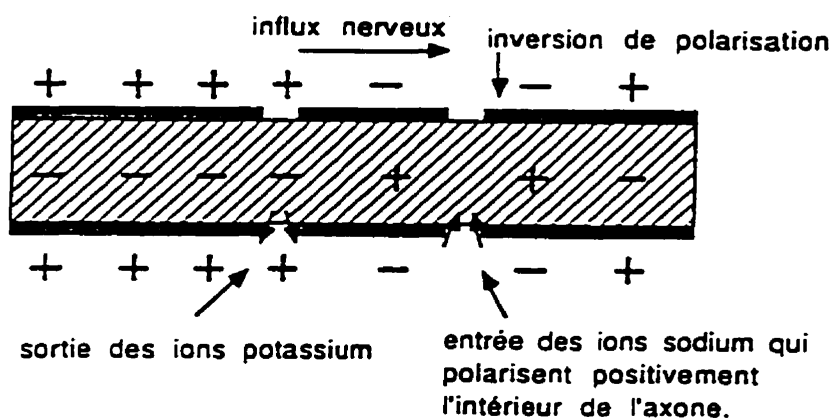
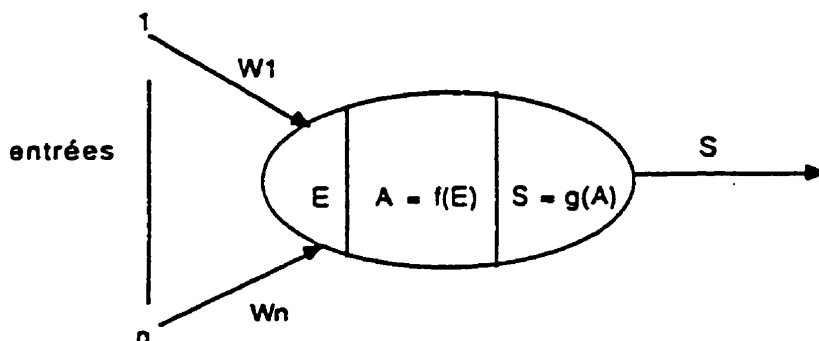


Figure 3.3: Propagation de l'influx nerveux dans le neurone biologique

## 3.2 FONDEMENT SUR LES NEURONES ARTIFICIELS

Un neurone artificiel fait une somme pondérée des potentiels d'actions qui lui parviennent et s'active suivant la valeur de cette sommation pondérée (figure 3.4). Si cette somme dépasse un certain seuil, le neurone est activé et transmet une réponse dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne transmet rien. Nous pouvons, à partir du fonctionnement d'un neurone artificiel, imaginer ce que fera un réseau de neurones artificiels. En effet, un réseau de neurones artificiels est une structure de traitement d'informations parallèlement distribuée, constituée d'éléments de traitement (ET) inter-reliés via les canaux de communication unidirectionnels (Brault, 1994). Chaque ET peut posséder une mémoire locale et effectuer un traitement arbitraire mais local. En effet, le traitement effectué par l'élément de traitement doit essentiellement dépendre de la valeur actuelle de ses





**Figure 3.4:** Représentation d'un neurone artificiel

signaux d'entrée et des valeurs emmagasinées dans sa mémoire locale. Un élément de traitement ne possède qu'une sortie qui peut être distribuée (copiée) à autant d'éléments de traitement que nécessaire. Le signal de sortie peut être d'un type mathématique quelconque.

### 3.2.1 Représentation

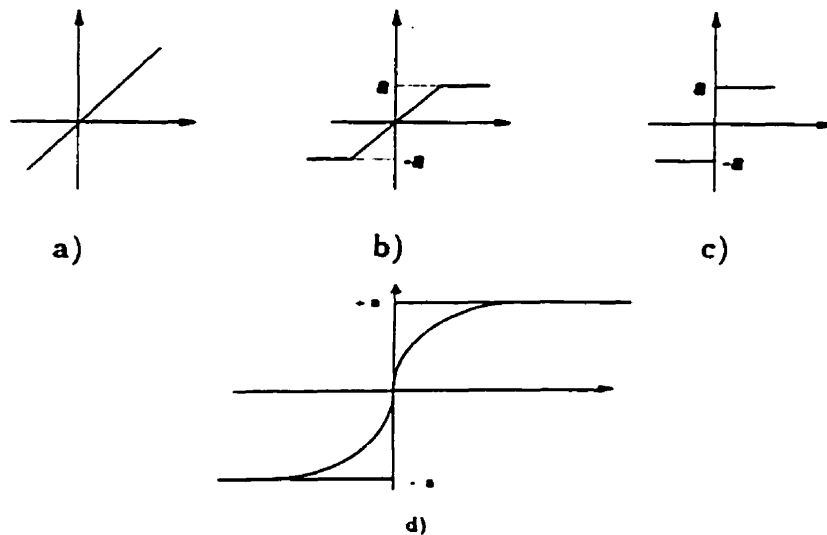
Cette représentation fait ressortir l'analogie que nous avons faite entre le neurone biologique et le neurone artificiel.

- **Les entrées**

Les entrées  $x_i$  issues d'autres neurones ou générées d'une manière quelconque, peuvent être binaires ou réelles.

- **La fonction d'entrée totale**

La fonction d'entrée totale définit le prétraitement effectué sur les entrées. Elle peut être booléenne, linéaire ou affine. Les types linéaire et affine sont les plus utilisés dans la conception des réseaux de neurones artificiels.



**Figure 3.5:** Fonctions d'activation: a) linéaire; b) linéaire avec saturation; c) fonction tout ou rien; d) fonction sigmoïde

- **La fonction d'activation**

On utilise généralement cinq types de fonctions d'activation.

- Linéaire (figure 3.5 a))
- Linéaire avec saturation (figure 3.5 b))
- Fonction tout ou rien (figure 3.5 c))
- Fonction sigmoïde (figure 3.5 d))
- Fonction stochastique

$$f(x) = \begin{cases} 1 & \text{avec probabilité } \frac{1}{1+\exp(\frac{-x}{T})} \\ 0 & \text{avec probabilité } 1 - \frac{1}{1+\exp(\frac{-x}{T})} \end{cases} \quad (3.1)$$

$T$  est un paramètre appelé température. Quand il tend vers 0, cette fonction tend vers la fonction seuil.

- On peut par ailleurs utiliser toute autre fonction croissante et impaire.

- **Fonction de sortie**

La fonction de sortie n'est autre que la fonction activation.

D'après la figure 3.4 et en supposant que la fonction d'entrée totale est de type linéaire, on peut écrire:

$$I_i = \sum_{j=1}^N w_{ij} x_j \quad (3.2)$$

où  $N$  est le nombre d'entrées et  $w_{ij}$  le poids pondéré entre les neurones  $j$  et  $i$ . Il faut signaler ici que  $j$  peut être une entrée quelconque ou une sortie d'un neurone.  $I_i$  est souvent appelé entrée nette. Si la fonction d'entrée totale est de type affine (ce qui est très souvent le cas), on a:

$$I_i = \sum_{j=1}^N w_{ij} x_j - b_i \quad (3.3)$$

$b_i$  est le seuil (ou le biais) du neurone  $i$ . La sortie  $y_i$  du neurone  $i$  est donnée par:

$$y_i = f\left[\sum_{j=1}^N w_{ij} x_j - b_i\right] \quad (3.4)$$

### 3.2.2 Topologie

Il existe principalement deux types de réseaux de neurones artificiels: des réseaux à couches et des réseaux entièrement connectés.

#### 3.2.2.1 Réseaux à couches

Cette configuration apparaît à la figure 3.6. Les neurones d'une même couche ne sont pas interconnectés. Les connexions existent entre une couche et celle qui la succède ou la précède. Ces connexions représentent les synapses auxquelles on a fait allusion dans le cas du neurone biologique. Pour modéliser ce rôle des synapses, on a choisi d'affecter à chaque connexion entre deux neurones artificiels, un poids qui pondère le signal transmis. Si les poids d'une connexion sont positifs, la connexion est excitatrice; sinon, elle est inhibitrice.

On trouve par ailleurs trois sortes de couches:

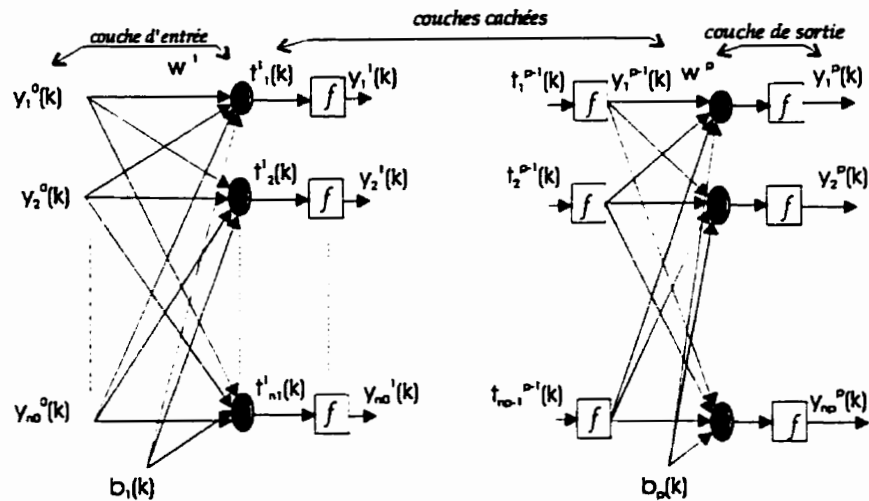


Figure 3.6: Réseaux multi-couches

- une couche d'entrée
- une couche ou plusieurs couches cachées
- une couche de sortie

La couche d'entrée reçoit l'information de l'extérieur, mais ne fait aucun traitement.

Les couches cachées servent de trait-d'union entre la couche d'entrée et celle de sortie.

La couche de sortie livre l'information à l'extérieur.

### 3.2.2.2 Réseaux entièrement connectés

Dans ce type de réseaux, chaque neurone est relié à tous les autres et possède un retour sur lui-même. Ce type de réseaux prend souvent le nom de réseaux récurrents (figure 3.7).

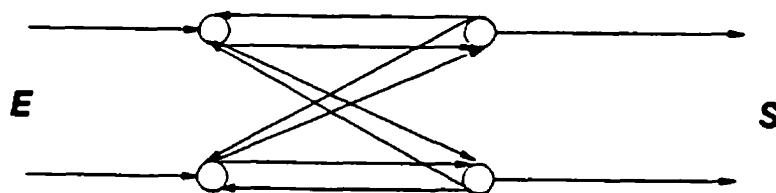


Figure 3.7: Réseaux connectés

### 3.2.3 Propagation des signaux

La propagation des signaux entre les neurones se fait généralement de l'entrée vers la sortie. Si cette propagation est à sens unique (entrée-sortie) on parle de réseaux à propagation avant (feedforward). Si elle se fait dans les deux sens, le réseau est dit à rétroaction (feedback). Ce deuxième cas est très souvent rencontré dans les réseaux récurrents et dans certains types d'apprentissage.

Tout réseau de neurones artificiels a une représentation distribuée des connaissances; les opérations sont locales et le traitement est non linéaire. Les opérations sont locales parce que le neurone ne voit que son entrée et sa mémoire locale pour générer une sortie.

## 3.3 DIFFÉRENTS MODÈLES DE RÉSEAUX NEURONAUX ARTIFICIELS

Il est plutôt difficile, aujourd'hui, de classer les différents réseaux. Chaque auteur propose une classification basée sur des critères subjectifs préalablement définis. C'est ainsi que certains auteurs classent les réseaux suivant leur mode d'apprentissage, d'autres les regroupent en fonction de leur topologie

d'interconnexion des éléments de traitement. R. Hecht-Nielsen (1990) les classe en trois modèles différents:

- Les réseaux de type associatif tels que les réseaux à propagation avant (feed-forward); les réseaux récurrents; l'associateur linéaire et le réseau "Learnmatrix".
- Les réseaux de type "Mapping" tels que le réseau basé sur les déclarations et le réseau "Prototype".
- Les réseaux de type spatio-temporel, stochastique et hiérarchique.

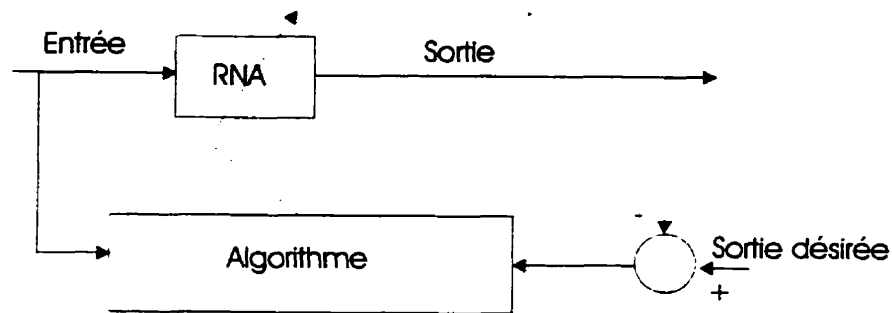
K. Mehrotra et *al.* (1997) les classent suivant leur architecture et distinguent:

- Les réseaux entièrement connectés;
- les réseaux à couches;
- les réseaux acycliques;
- les réseaux à propagation avant;
- les réseaux modulaires.

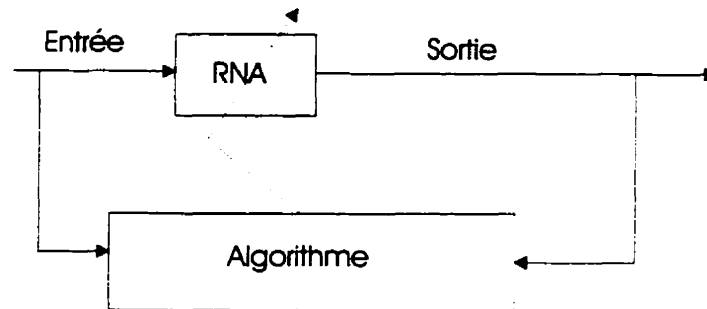
Cette classification se rapproche de celle de R.Hecht-Nielsen.

### 3.4 APPRENTISSAGE DES RÉSEAUX

Les réseaux de neurones artificiels prennent de plus en plus d'ampleur à cause de leur capacité d'apprentissage. On rencontre à l'heure actuelle plusieurs lois d'apprentissage qui diffèrent, les unes des autres, par leur formalisme. De ces lois, quatre grandes classes sont les plus utilisées.



**Figure 3.8:** Apprentissage supervisé



**Figure 3.9:** Apprentissage non supervisé

### 3.4.1 Apprentissage supervisé

Ce type d'apprentissage est aussi appelé apprentissage actif. Les valeurs de l'entrée et de la sortie désirée sont fournies au réseau qui, à son tour, ajuste les poids en tenant compte de l'entrée et de l'erreur de sortie (figure 3.8).

### 3.4.2 Apprentissage non supervisé

L'Apprentissage non supervisé est aussi appelé apprentissage auto-organisateur. Seules les valeurs d'entrée sont fournies au réseau. Le réseau ajuste

les poids en tenant compte de l'entrée et de la sortie calculée et non de l'erreur (figure 3.9).

### 3.4.3 Apprentissage renforcé

Le terme *renforcé* vient des études menées sur le dressage des animaux en psychologie expérimentale où on s'intéresse à l'accomplissement d'un événement et la relation propre qui tend à accroître la probabilité que cet événement se repète dans les mêmes conditions ( Kimble, 1961). Le terme apprentissage renforcé qui n'est pas utilisé par les psychologues mais par les théoriciens en ingénierie et en intelligence artificielle est l'ensemble des algorithmes d'apprentissage basés sur ce principe de renforcement.

### 3.4.4 Apprentissage par rétropropagation d'erreur

La méthode de rétropropagation d'erreur minimise l'erreur issue de l'approximation du réseau en trouvant les poids optimaux. Beaucoup de recherches ont montré que cette méthode calcule mieux les poids optimaux surtout dans le cas des réseaux à couches. Lapedes et Farber (1987) ont largement contribué à l'analyse théorique et pratique de cette méthode.

#### 3.4.4.1 Méthode de la descente du gradient

Dans les réseaux multi-couches, il se pose le problème d'ajustage des poids des neurones des couches cachées. Un réseau de neurones à  $p + 1$  couches, d'entrée  $y_j^0$  ( $j = 1, \dots, n_0$ ) et de sortie  $y_j^p$  ( $j = 1, \dots, n_p$ ) peut être représenté tel que montré à



la figure 3.6. Sous cette représentation, il est régi par l'équation suivante:

$$y_i^p = f_i^p\left(\sum_{j=1}^{n_{p-1}} w_{ij}^p y_j^{p-1} + b_i\right) \quad (3.5)$$

où  $w^p$  est la matrice de poids ( $n_p \times n_{p-1}$  associés à la  $p^{eme}$  couche. Le vecteur  $b_i (i = 1, 2 \dots p)$  représente la valeur du biais pour chaque neurone  $i$ . La fonction  $f_i^p[.]$  est un opérateur non décroissant pouvant être linéaire ou non linéaire appliqué à chaque sortie du neurone  $i$ . La fin de l'apprentissage est dictée par la convergence de la matrice des poids  $w^p$  vers des valeurs constantes. L'on atteint cet objectif en minimisant l'erreur quadratique entre les sorties  $y_i^p$  données par le réseau et les sorties désirées  $y_d$  pour un vecteur d'entrées donné. Il s'agit donc de minimiser la fonction coût ci-dessous:

$$J = \frac{1}{2} \sum_{k=1}^{n_e} \sum_{s=1}^{n_p} e_s^2(k) \quad (3.6)$$

avec  $e_s(k) = y_d(k) - y_i^p(k)$ ;  $n_e$  le nombre d'échantillons et  $n_p$  la longueur du vecteur de sortie  $y$ .

Pour minimiser la fonction coût  $J$ , plusieurs méthodes sont exploitées dans la littérature. La plus populaire étant la méthode du gradient appliquée à la rétropropagation de l'erreur. L'idée principale de cette méthode d'apprentissage est de propager l'erreur de sortie sur les neurones des couches cachées à partir de l'erreur de la couche de sortie. Le gradient calculé à la  $q^{ieme}$  couche de  $n_q$  neurones est donné par:

$$g_{ij}^q = \frac{\partial J}{\partial w_{ij}^q} = \sum_{k=1}^{n_e} \sum_{s=1}^{n_p} \frac{\partial J}{\partial y_s(k)} \frac{\partial y_s(k)}{\partial w_{ij}^q} \quad (3.7)$$

qui peut se réduire à:

$$g_{ij}^q = - \sum_{k=1}^{n_e} \sum_{s=1}^{n_p} e_s(k) \frac{\partial y_s(k)}{\partial w_{ij}^q} \quad (3.8)$$

Compte tenu de l'expression de  $y^s$ , la relation ci-dessus devient (M. Saad et al, 1994):

$$g_{ij}^q = - \sum_{k=1}^{n_e} \sum_{s=1}^{n_p} e_s(k) \sum_{r=1}^{n_q} \frac{\partial y^s(k)}{\partial y_r^q(k)} \frac{\partial y_r^q(k)}{\partial w_{ij}^q} \quad (3.9)$$

Pour  $q = p$  (couche de sortie)

$$\begin{aligned} \frac{\partial y_r^q(k)}{\partial w_{ij}^q(k)} &= \frac{\partial y_r^q(k)}{\partial t_i^q(k)} \frac{\partial t_i^q(k)}{\partial w_{ij}^q(k)} \\ &= \dot{f}_q[t_i^q(k)] y_j^{q-1} \delta(r-i) \end{aligned} \quad (3.10)$$

où  $\dot{f}_q[\cdot]$  est la dérivée de la fonction  $f_q[\cdot]$  et  $\delta(r-i)$  représente l'opérateur de Kronecker tel que:

$$\delta(r-i) = \begin{cases} 1 & \text{si } r=i \\ 0 & \text{si } r \neq i \end{cases} \quad (3.11)$$

Les équations (3.9) et (3.10) permettent d'écrire l'équation (3.8) comme ci-dessous:

$$g_{ij}^q = \frac{\partial J}{\partial w_{ij}^q} = - \sum_{k=1}^{n_e} \rho_i^q(k) y_j^{q-1}(k) \quad (3.12)$$

avec:

$$\rho_i^q = \dot{f}_q[t_i^q(k)] e_i(k) \quad \text{si } q = p \quad (3.13)$$

et

$$\rho_i^q = \dot{f}_q[t_i^q(k)] \sum_{s=1}^{n_q} \rho_s^{q+1}(k) w_{si}^{q+1} \quad \text{si } q \neq p \quad (3.14)$$

l'équation d'adaptation des poids est donnée par:

$$w_{ij}^q(k+1) = w_{ij}^q(k) - \eta g_{ij}^q \quad (3.15)$$

où  $\eta$  est le coefficient d'apprentissage ajustable à chaque itération;  $g_{ij}^q$  ayant été déterminé par les précédents calculs.

Outre cette méthode de rétropropagation de l'erreur, d'autres méthodes sont utilisées; c'est le cas des méthodes de Newton et Quasi-Newton.

### 3.4.4.2 Méthode de Newton

Elle est basée sur une approximation de la fonction en présence où l'on ne tient compte que des premiers termes de la série de Taylor. La série considérée est tronquée après trois termes de sorte qu'elle puisse s'écrire:

$$f(x) = f(x_k) + \nabla f(x_k)^t \Delta x + \frac{1}{2} \Delta x^t \nabla^2 f(x_k) \Delta x + o(\Delta x^3) \quad (3.16)$$

Posons que:

$$\tilde{f}(x; x_k) = f(x_k) + \nabla f(x_k)^t \Delta x + \frac{1}{2} \Delta x^t \nabla^2 f(x_k) \Delta x \quad (3.17)$$

où  $\tilde{f}(x; x_k)$  est une fonction approximative construite à  $x_k$ . Utilisons cette fonction approximative de  $f(x)$  pour générer une séquence itérative forçant  $x_{k+1}$ , le point suivant de la séquence, d'être un point où le gradient de  $\tilde{f}(x; x_k)$  s'annule.

$$\nabla \tilde{f}(x; x_k) = \nabla f(x_k) + \nabla^2 f(x_k) \Delta x = 0 \quad (3.18)$$

soit

$$\Delta x = -\nabla^2 f(x_k)^{-1} \nabla f(x_k) \quad (3.19)$$

sachant que:

$$x_{k+1} = x_k + \Delta x \quad (3.20)$$

On pose facilement

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k) \quad (3.21)$$

C'est sur cette approche qu'est basée l'adaptation des poids dans le cas de l'apprentissage par rétropropagation de l'erreur. On peut voir que la descente du gradient est bâtie sur cette approche.

### 3.4.4.3 Autres méthodes

**Méthode Quasi-Newton.** Appelée aussi méthode à métrique variable, elle est une approximation de l'algorithme de Newton (Fletcher et Powell, 1963). Le principe consiste à estimer l'hessien inverse à chaque itération.

Brayden-Fletcher-Goldfarb-Shanno (BFGS) (Fletcher, 1987) montrent que  $H(k)$  peut être estimée à chaque itération par la relation ci-dessous:

$$H(k+1) = H(k) + \left[1 + \frac{\Delta^t H(k) \Delta}{A^t \Delta}\right] \frac{A A^t}{A^t \Delta} - \frac{A \Delta^t H(k) + H(k) \Delta A^t}{A^t \Delta} \quad (3.22)$$

où  $\Delta$  est la différence des gradients  $\nabla(k+1)$  et  $\nabla(k)$  puis  $A$  est donnée par:

$$A = -\eta H(k) \nabla(k) \quad (3.23)$$

**Méthode de Levenberg Marquardt.** Une autre méthode d'apprentissage est la méthode de rétropropagation basée sur l'approximation de Levenberg Marquardt. L'équation d'adaptation des poids est donnée par:

$$w(k+1) = w(k) + (J_b^t J_b + \mu I)^{-1} J_b^t e \quad (3.24)$$

où  $J_b$  est la matrice jacobienne (dérivées de chaque erreur d'estimation par rapport aux poids);  $e$  est le vecteur d'erreur et  $\mu$  est un scalaire ajustable en fonction du vecteur d'erreur  $e$ . Plus l'erreur est petite, plus grand est  $\mu$ . Cette méthode est programmée dans un sous-programme par Mathworks Matlab dans "Neural Network Toolbox" (Demuth et Beale, 1994). Le nom d'appel du sous-programme est "trainlm".

## Conclusion

Ce chapitre a permis de faire un survol rapide sur les notions élémentaires mais nécessaires dans la compréhension du fonctionnement des réseaux de neurones artificiels. C'est ainsi que nous avons passé en revue quelques types de réseaux neuronaux et certaines méthodes d'apprentissage très souvent utilisées dans l'analyse des réseaux de neurones artificiels.

Bien que n'étant pas plastique (la plasticité étant cette possibilité qu'ont certains réseaux de ne pas "oublier" les anciennes informations apprises lors de l'apprentissage de nouvelles), le réseau rétropropagation est le plus utilisé. Cette préférence est liée à ses capacités de mémorisation de l'information dans le temps et dans l'espace. Il est, par ailleurs, facile d'implantation.

Il faut noter que l'étude des réseaux neuronaux est assez vaste et très largement expliquée dans la littérature. Nous avons présenté ici l'essentiel qui nous servira dans la suite de la recherche. Les théorèmes et les paramètres importants qui interviendront, de temps à autre dans le cadre de cette étude, seront expliqués et présentés chaque fois que le besoin se fera sentir.

## Chapitre 4

# CONCEPTION DU FILTRE NEURONAL BASÉ SUR LA STRUCTURE DU FILTRE DE KALMAN

### Introduction

Avant d'analyser la structure du filtre neuronal, ce chapitre donne quelques applications des réseaux neuronaux artificiels en automatique dans le but d'élucider l'utilisation non moins importante des réseaux neuronaux dans le domaine de la commande.

L'analyse du filtre de Kalman démontre que pour avoir la matrice de gain optimal, il faut minimiser la variance de l'erreur d'estimation. Ce postulat permet de trouver les expressions récursives ci-dessous:

$$\left\{ \begin{array}{l} x_{k+1|k+1} = F_k x_{k|k} + B_k u_k + K_{k+1} (y_{k+1} - C_k F_k x_{k|k} + B_k u_k) \\ \Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \Sigma_{k+1|k} C_k^t [C_k \Sigma_{k+1|k} C_k^t + R_k]^{-1} C_k \Sigma_{k+1|k} \\ K_{k+1} = \Sigma_{k+1|k} C_k^t [C_k \Sigma_{k+1|k} C_k^t + R_k]^{-1} \\ \Sigma_{k+1|k+1} = (I - K_{k+1} C_k) \Sigma_{k+1|k} \\ x_{k+1|k+1} = x_{k+1|k} + K_{k+1} (y_{k+1} - C_k x_{k+1|k}) \end{array} \right. \quad (4.1)$$

qui donnent, entre autres, le gain  $K_{k+1}$  en fonction des statistiques des bruits. En gardant la même structure que le filtre de Kalman, nous remplaçons son gain par un réseau de neurones artificiels qui agira sur ses poids afin de minimiser l'erreur à l'entrée.

## 4.1 LES RÉSEAUX NEURONAUX ARTIFICIELS DANS LA COMMANDE

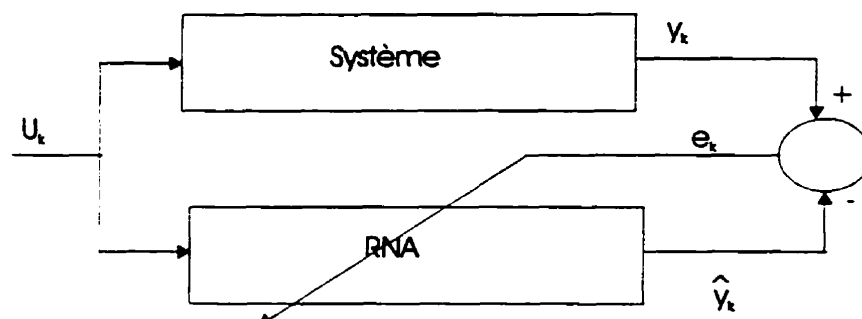
Plusieurs études ont été menées par différents auteurs où les réseaux de neurones artificiels (RNA) sont utilisés soit comme estimateurs soit comme contrôleurs. A titre d'exemple, on peut citer les travaux de : Guez (1987), Yoshiaki et Toshiyuki (1992), Kuschewski et *al.*(1993), Bavarian (1988), Levin (1991), Bulsari (1993). Narendra (1990, 1991, 1994) est l'un des chercheurs les plus indiqués dans le domaine de la commande par réseaux de neurones.

### 4.1.1 Réseaux de neurones comme estimateurs

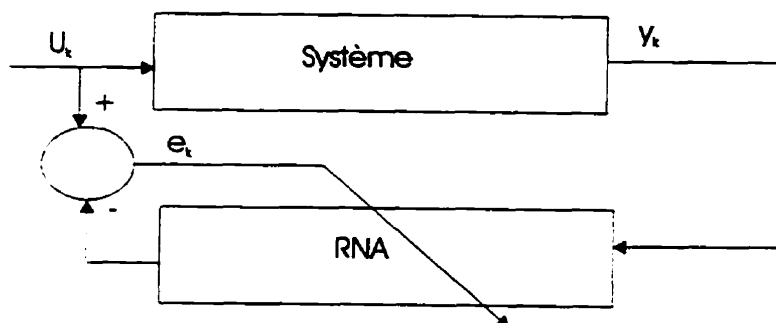
Les réseaux sont très souvent utilisés pour identifier les fonctions de transfert des systèmes dynamiques dans le but de déterminer leurs paramètres. Le réseau apprend le comportement dynamique du système. Cet apprentissage peut se faire de deux façons différentes.

#### 4.1.1.1 Apprentissage de la dynamique directe du système

La figure 4.1 montre le cas d'apprentissage de la dynamique directe du système. Le problème d'identification consiste à estimer les paramètres du modèle de manière à optimiser un critère de performance préalablement défini et basé sur l'erreur entre la sortie du système réel et la sortie du système identifié. L'erreur  $e_k$  permet



**Figure 4.1:** Dynamique directe



**Figure 4.2:** Dynamique inverse

d'ajuster les poids du réseau pour que les deux sorties  $\hat{y}_k$  et  $y_k$  soient identiques.

#### 4.1.1.2 Apprentissage de la dynamique inverse du système

Une autre approche d'estimation des paramètres du système est l'utilisation de l'apprentissage de la dynamique inverse comme illustré à la figure 4.2.

A la convergence du réseau, sa dynamique est l'inverse de celle du système.

Dans les deux cas de la dynamique inverse ou directe, le réseau est un émulateur.



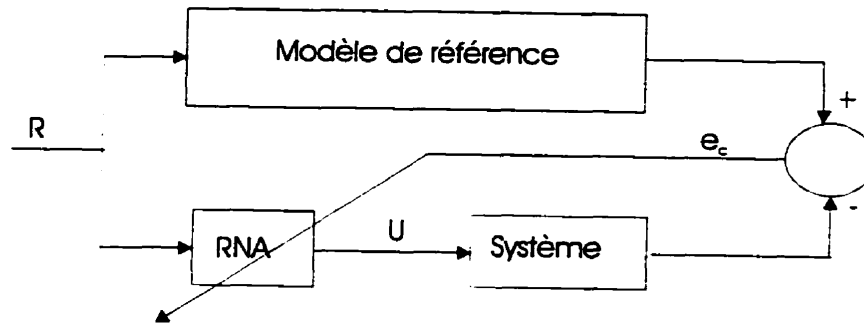


Figure 4.3: Commande adaptative directe

### 4.1.2 Réseau de neurones comme contrôleurs

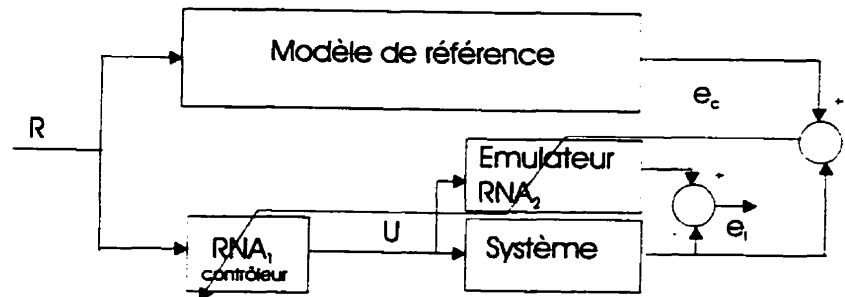
Un des types de commande où interviennent le plus souvent les réseaux est la commande adaptative qui apparaît sous deux configurations:

#### 4.1.2.1 Commande adaptative directe

La figure 4.3 montre le schéma d'une commande adaptative directe par réseau de neurones. L'implantation de cette méthode pose des problèmes pratiques dans la mesure où l'erreur  $e_k$  se trouve entre le système et le contrôleur (RNA). On n'a pas accès à l'erreur de sortie du réseau, ce qui rend difficile l'ajustage des poids. Pour contourner ce problème, on utilise la commande adaptative indirecte.

#### 4.1.2.2 Commande adaptative indirecte

La figure 4.4 montre cette approche. Le réseau  $RNA_2$  est un émulateur qui apprend la dynamique du système. A la fin de cet apprentissage, le réseau  $RNA_2$  se comporte comme le système. Il est donc possible de rétropropager l'erreur de sortie  $e_c$  à travers  $RNA_2$  pour l'avoir à la sortie du contrôleur  $RNA_1$ . De cette manière on pourra



**Figure 4.4:** Commande adaptative indirecte

procéder à l'apprentissage du contrôleur à partir de l'erreur de sortie du système  $e_c$ . Beaucoup de chercheurs tels que Nguyen et Widrow (1990), Psaltis (1988) ont utilisé cette approche pour construire des contrôleurs neuronaux. Par contre Xianzhong et Kang (1993) ont montré qu'il est possible de commander les systèmes en utilisant l'approche directe. Le principe est basé sur la définition de la direction du système. En effet, si la sortie  $y(t)$  d'un système croît (ou décroît) de façon monotone lorsque la commande  $u(t)$  croît (ou décroît), le système est dit répondre positivement (ou négativement). Dans les deux cas le système a une réponse monotone. Pour un système monovariante,  $y(t) = G(u(t))$ , si le système répond positivement (ou négativement), la direction du système est  $D(G) = 1$  (ou  $D(G) = -1$ ). Les auteurs montrent qu'il est à cet effet possible d'entraîner le contrôleur à partir de l'erreur de sortie du système en utilisant l'apprentissage par rétropropagation.

### 4.1.3 Identification des paramètres par les réseaux neuronaux

Beaucoup de travaux faisant intervenir les réseaux neuronaux dans l'identification des paramètres des systèmes dynamiques ont été réalisés. Raol (1995) utilise le réseau récurrent de Hopfield pour bâtir un algorithme d'estimation des paramètres des systèmes dynamiques. Ruchti et *al.* (1992) ont mis sur pied un processus d'apprentissage de réseaux pour identifier les paramètres des systèmes non linéaires. L'identification des paramètres des modèles *ARMA* par les réseaux neuronaux est aussi très développée. C'est surtout dans le domaine hydraulique que cette identification est le plus utilisée. Markus (1995) dans sa thèse de doctorat, a procédé à la prédiction des débits par réseaux neuronaux et a comparé ses résultats avec ceux donnés par le modèle de la régression linéaire, le modèle de corrélation canonique et le modèle *ARMA*. Les résultats donnés par ces prédicteurs étaient très semblables.

### 4.1.4 Autres types de commande par réseaux

Certains auteurs, comme Xu et Vukovich (1993), combinent les réseaux neuronaux à la logique floue. Ce type de commande est très souvent connu dans le domaine des réseaux comme la commande mixte. Les réseaux de neurones sont aussi très utilisés pour ajuster les paramètres d'un contrôleur classique. Kumar et Guez (1991) utilisent un réseau ART pour faire le placement des pôles d'un système. Chen (1990) ajuste les paramètres d'un contrôleur adaptatif.

Les plus grands problèmes qui surgissent dans la commande des systèmes par réseaux de neurones sont liés au fait que, fondamentalement, les réseaux sont conçus pour être statiques. Or les systèmes à commander ont une dynamique; il faut donc

utiliser les apprentissages qui tiennent compte de cette dynamique. Par ailleurs, le problème de plasticité se pose avec acuité. Pour trouver une solution à ce problème, Carpenter et Grossberg (1987) ont mis sur pied le réseau ART (Adaptive Resonance Theory). On l'utilise surtout pour la reconnaissance de formes mais très rarement pour la commande. Ce réseau a une faiblesse qui est sa très grande instabilité et ses difficultés d'implantation.

## 4.2 LE FILTRE NEURONAL

Le modèle de base du filtre neuronal que nous proposons ici est le reflet de la structure du filtre de Kalman où son gain est remplacé par un réseau de neurones artificiels. La structure du nouveau filtre est celle apparaissant à la figure 4.5.

### 4.2.1 L'estimateur neuronal

La structure de l'estimateur neuronal est exactement celle de la figure 4.5. L'apprentissage du réseau se fait généralement lorsqu'on possède le vecteur d'informations entrée/sortie où la sortie représente la valeur désirée. La structure de la figure 4.5 ne nous permet pas d'avoir, dès le départ, les sorties désirées. Aussi est-il nécessaire d'utiliser le réseau comme un système adaptatif, c'est-à-dire un apprentissage en ligne compte tenu de l'erreur. Ceci ne permettra pas d'obtenir une convergence des poids mais des poids qui évolueront en fonction de l'erreur d'estimation. Dans ce cas d'ajustement des poids à chaque itération, la fonction objective de l'équation (3.6) se réduit à la forme suivante:

$$J(k) = 1/2 e^t(k) e(k) \quad (4.2)$$

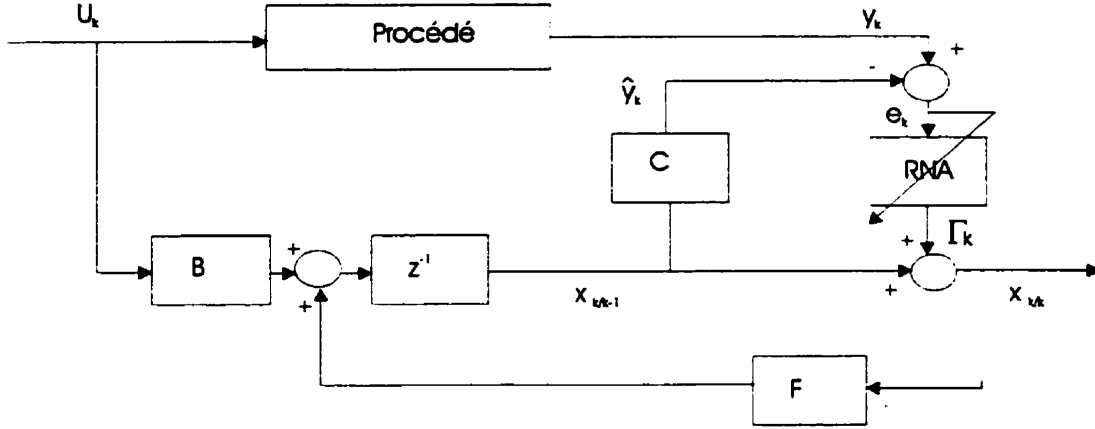


Figure 4.5: Structure du filtre neuronal

où  $e(k) = (y_k - \hat{y}_k)$ .

Ecrivons la variation du critère  $J(k)$  par rapport aux poids du réseau:

$$\frac{\partial J(k)}{\partial w_{ij}} = -(y_k - \hat{y}_k)^t \frac{\partial \hat{y}_k}{\partial w_{ij}} \quad (4.3)$$

$y_k$  étant indépendant des poids du réseau.

L'apprentissage par la méthode de descente du gradient permet de poser:

$$\Delta w_{ij} = -\eta e^t(k) \frac{\partial \hat{y}_k}{\partial w_{ij}} \quad (4.4)$$

où  $\eta$  est le coefficient d'apprentissage généralement  $\ll 1$ . S'il est possible de calculer la variation de la sortie estimée par rapport au poids ( $\frac{\partial \hat{y}_k}{\partial w_{ij}}$ ), nous pouvons évidemment déterminer  $\Delta w_{ij}$  qui permet d'ajuster optimalement les poids.

#### 4.2.2 Calcul de la variation de la sortie estimée par rapport au vecteur des poids ( $\frac{\partial \hat{y}_k}{\partial w_{ij}}$ )

À partir de la figure 4.5, on peut écrire:

$$\hat{y}_k = C x_{k/k-1} \quad (4.5)$$

et

$$x_{k/k-1} = Fx_{k-1/k-1} + Bu_{k-1} \quad (4.6)$$

en outre

$$x_{k-1/k-1} = Fx_{k-2/k-2} + Bu_{k-2} + \Gamma_{k-1} \quad (4.7)$$

d'où

$$\frac{\partial \hat{y}_k}{\partial w_{ij}} = CF^2 \frac{\partial}{\partial w_{ij}} x_{k-2/k-2} + CF \frac{\partial}{\partial w_{ij}} \Gamma_{k-1} \quad (4.8)$$

en posant

$$x_{k/k} = \hat{x}_k \quad (4.9)$$

l'équation(4.8) s'écrit:

$$\frac{\partial \hat{y}_k}{\partial w_{ij}} = CF^2 \frac{\partial \hat{x}_{k-2}}{\partial w_{ij}} + CF \frac{\partial \Gamma_{k-1}}{\partial w_{ij}} \quad (4.10)$$

#### 4.2.3 Calcul de la variation de la sortie du réseau par rapport au vecteur des poids $(\frac{\partial \Gamma_{k-1}}{\partial w_{ij}})$

Si la structure du réseau est connue ainsi que les fonctions d'activation, le calcul de cette variation se fait par la rétropropagation de l'erreur d'estimation basée sur la descente du gradient. Cette procédure de calcul a été développée au paragraphe 3.4.4.1 du précédent chapitre. Il suffit, en somme, de remplacer  $y_k$  par  $\Gamma_k$  dans les équations.

#### 4.2.4 Calcul de la variation des états estimés par rapport au vecteur des poids $(\frac{\partial \hat{x}_k}{\partial w_{ij}})$

Soit un réseau de  $p$  couches de  $n_p$  neurones chacune et  $n_r$  entrées; soit un état estimé  $\hat{x}_i(k)$ ; le calcul de la variation  $\frac{\partial \hat{x}_i(k)}{\partial w_{ij}}$  se fait comme décrit à la figure 4.6. Narendra

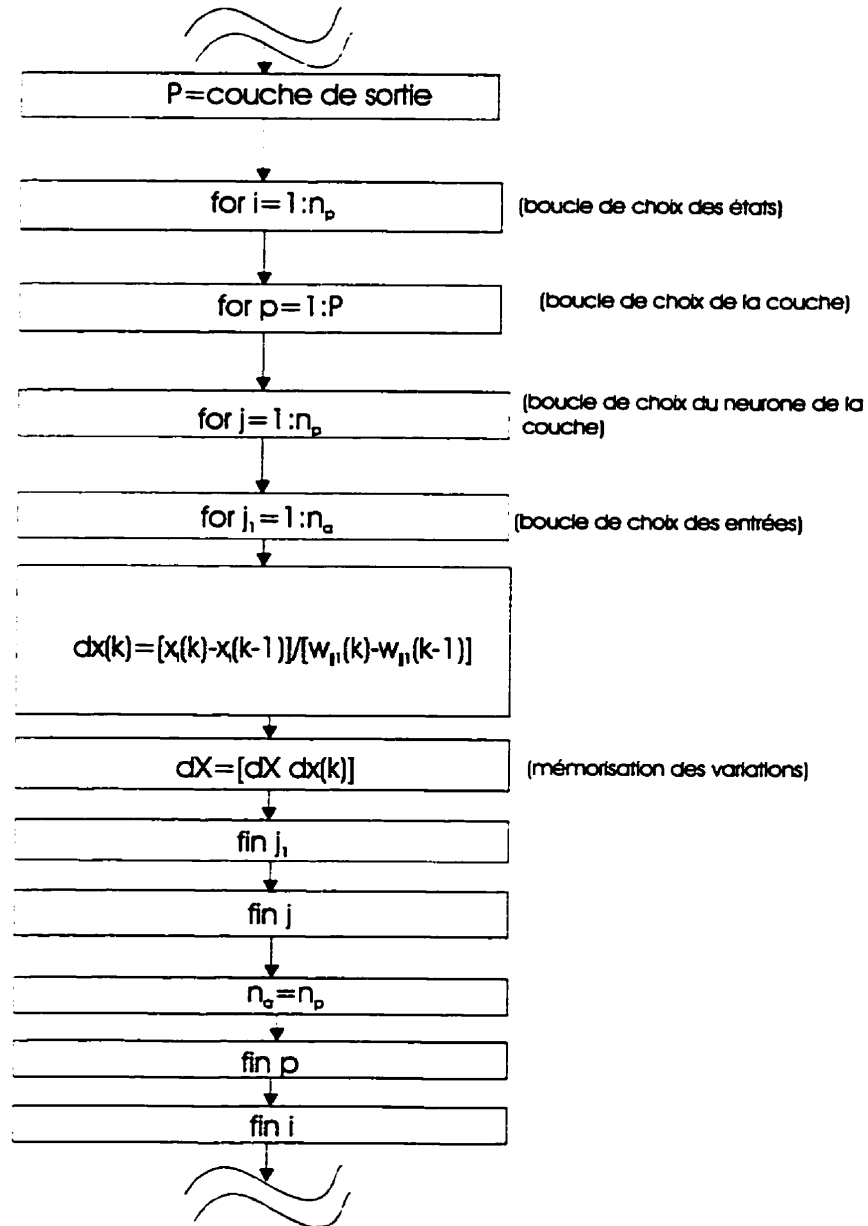


Figure 4.6: Organigramme de calcul de  $\frac{\partial \hat{x}_i(k)}{\partial w}$  à l'instant  $k$ .

et Parthasarathy (1990) proposent la même approche.

En effet, la loi d'apprentissage est celle donnée ci-dessous:

$$w_{ij}(k+1) = w_{ij}(k) + \eta e^t(k) CF \left[ F \frac{\partial \hat{x}_{k-2}}{\partial w_{ij}} + \frac{\partial \Gamma_{k-1}}{\partial w_{ij}} \right] \quad (4.11)$$

La connaissance de cette loi d'apprentissage permet de calculer les poids optimaux qui minimisent l'erreur d'estimation. Dans le cas d'un apprentissage hors ligne où les poids doivent converger, Jacobs (1988) dit que cette convergence est en grande partie fonction du coefficient d'apprentissage  $\eta$ . Il propose, entre autres, de choisir un coefficient d'apprentissage pour chaque poids, ce qui alourdit les calculs et suggère en outre que le coefficient d'apprentissage varie dans le temps. Il conseille enfin que si le coefficient d'apprentissage pour un poids donné a un signe qui alterne à chaque itération, il devra être décrémente. Par ailleurs, si son signe est constant, il devra être incrémenté. Les auteurs (Kandil, Khorasani, Patel et Sood, 1993) trouvent le coefficient d'apprentissage par la relation ci-dessous:

$$\eta_k^* = \frac{e_k^t S_k e_k}{e_k^t S_k^2 e_k} \quad (4.12)$$

où  $e_k$  est l'erreur d'apprentissage et  $S_k$  une matrice définie positive donnée par:

$$S_k = \dot{f}(Net_k) H \dot{f}(Net_k) \quad (4.13)$$

$f$ : fonction d'activation

$Net_k$ : l'entrée nette du réseau

Ces suggestions heuristiques sont difficiles à implanter, aussi choisit-on souvent un coefficient d'apprentissage constant tel que  $0 < \eta < 1$ .

L'erreur d'estimation à rétropropager à travers les couches dans le calcul du gradient de la méthode de rétropropagation est à la sortie du réseau de neurones. Dans le filtre que nous proposons, l'erreur est plutôt à l'entrée du réseau. Ce fait, à première vue, peut sembler paradoxal. Il ne se pose pourtant aucun problème pour le développement des équations dans la mesure où la structure du filtre comporte une rétroaction de la sortie. On peut exprimer l'erreur comme une fonction de la



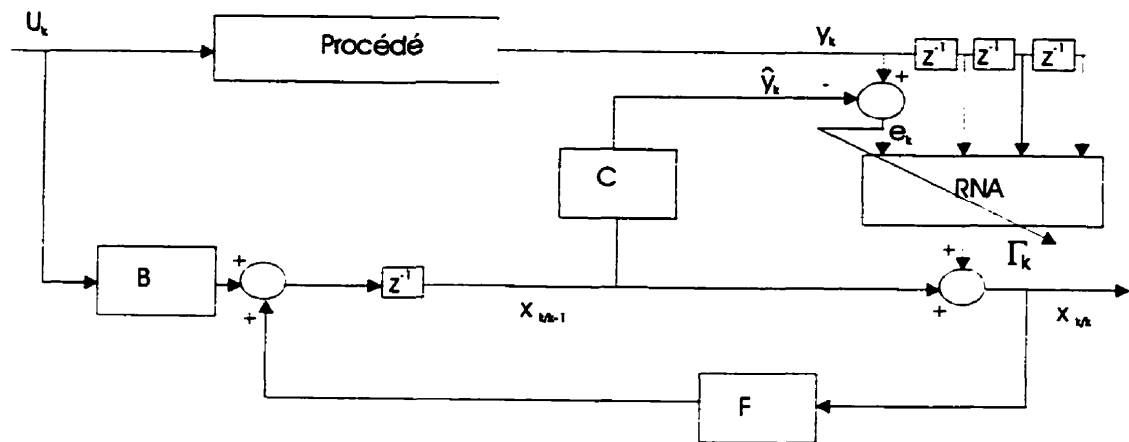


Figure 4.7: Structure du prédicteur neuronal

sortie du réseau, et le calcul de la variation de la sortie estimée par rapport au vecteur des poids ( $\frac{\partial \hat{y}_k}{\partial w_{i,j}}$ ) se fait comme présenté au paragraphe 4.2.

#### 4.2.5 Le prédicteur neuronal

La structure du prédicteur neuronal est un peu différente de la structure conventionnelle du filtre de Kalman. Comme le montre la figure 4.7, l'entrée du réseau n'est pas seulement l'erreur mais aussi des valeurs passées de la mesure  $y_k$ . Le but visé par l'ajout des valeurs passées de  $y_k$  est de permettre au réseau de minimiser l'erreur en tenant compte de la dynamique du système. Le réseau pourra accroître sa capacité de prédiction car son espace d'apprentissage est composé des valeurs réelles des mesures.

La loi d'apprentissage du prédicteur neuronal est la même que celle de l'estimateur seule la dimension du vecteur d'entrée change.

### 4.2.6 Choix du modèle du réseau

L'objectif du réseau est de minimiser l'erreur  $e_k$  de la figure 4.5 en agissant sur ses poids. Cette minimisation peut se faire soit par apprentissage en ligne ou par apprentissage hors ligne. Le réseau devra tenir compte du temps, de l'espace et avoir une capacité de mémorisation. Entre autres, le réseau à propagation avant permet de minimiser l'erreur  $e_k$  et possède les caractéristiques précédemment citées, d'où son choix dans cette étude.

### 4.2.7 Choix du nombre de couches du réseau neuronal

Fausette (1994) affirme qu'un réseau de neurones à une seule couche cachée et une couche de sortie est capable d'approximer n'importe quelle fonction. Aussi allons-nous utiliser dans cette recherche un réseau à une couche cachée et une couche de sortie.

Par ailleurs, il n'existe quasiment aucun calcul rigoureux du nombre de neurones à utiliser pour une couche cachée donnée d'un réseau de neurones. La plupart des chercheurs procèdent par expérimentation et retiennent la configuration qui aboutit à de bons résultats. Néanmoins, Swingler (1996) propose le calcul du nombre de neurones à utiliser sur une couche cachée à partir du nombre des entrées du réseau en se basant sur le théorème de Kolmogorov. Il suggère de calculer le nombre de neurones ( $n_o$ ) de la couche cachée par la relation:

$$n_o = 2n_a + 1 \quad (4.14)$$

où  $n_a$  est le nombre d'entrées.

## Conclusion

Le présent chapitre a permis de formuler le nouveau filtre neuronal et d'écrire ses équations caractéristiques. En effet, la structure de base du filtre neuronal peut être légèrement transformée pour obtenir celle du prédicteur neural. Les équations de base du filtre neuronal sont valables aussi bien pour l'estimateur que pour le prédicteur neuronal. On constate qu'en remplaçant le gain du filtre de Kalman par un réseau de neurones artificiels, il est théoriquement possible d'estimer les états des systèmes dynamiques. Aussi la validation de cette structure sera effective grâce à une vérification expérimentale. Le chapitre suivant définit le cadre, la nature et les bases de l'expérimentation.

## Chapitre 5

# SIMULATION DU FILTRE NEURONAL

### Introduction

Le présent chapitre traite de la simulation du filtre neuronal. On s'intéresse à l'aspect estimatif et prédictif du filtre. Dans le cas de l'estimation, il s'agit de déterminer les variables d'état correspondant au mouvement longitudinal d'un avion Boeing 747. Pour le filtre prédictif, on prendra le cas de la prédiction des apports naturels de trois réservoirs de stockage d'eau de l'Hydro-Québec.

Le meilleur estimateur dont on dispose de nos jours est le filtre de Kalman et l'optimalité de ses estimés est surtout liée à la connaissance quasi parfaite des statistiques des bruits du système.

Par ailleurs, le souci de mieux s'informer sur le futur a toujours hanté l'homme et l'a poussé à se doter des moyens susceptibles de le renseigner par rapport aux événements à venir. Ce besoin constant de connaître le futur se fait le plus souvent ressentir en commande où la nécessité d'avoir l'information sur le futur est primordiale dans le but de calculer la commande adéquate. Dans l'état actuel des connaissances, une fois de plus, le filtre de Kalman donne les meilleurs résultats

quant à la prédiction. Il est néanmoins limité dans sa formulation; l'estimé prédictif qu'il calcule étant juste à un pas dans le temps et nécessite la connaissance des statistiques des bruits du système. Des tentatives pour utiliser le filtre de Kalman comme prédicteur à  $m$  pas ( $m > 1$ ) ont été faites par certains auteurs tels que Kumar et Varaiya (1986); Davis et Varaiya (1973) qui dégagent les conditions d'optimalité nécessaires et suffisantes dans le cas des systèmes décrits par des équations différentielles stochastiques pour une commande basée sur des observations incomplètes. Pour avoir la prédiction à l'instant  $t_m$  lorsqu'on se situe à l'instant  $t_p$  ( $t_m > t_p$ ), les  $t_m - t_p - 1$  informations entre  $t_m$  et  $t_p$  étant inconnues, entraînent des covariances infinies à ces instants. Cette difficulté est contrée en supposant que le système (entre  $t_m$  et  $t_p$ ) est déterministe.

Les études basées sur ces hypothèses montrent que le filtre de Kalman dans ce contexte donne des résultats passablement bons. D'autres recherches se poursuivent toujours pour développer de nouvelles méthodes de prédiction à long terme (Sweigend and Gershenfeld, 1993). Par ailleurs, à partir des informations entrée/sortie d'un système quelconque, le RNA peut procéder à un apprentissage au bout duquel il acquerra la capacité de prédire les états du système à des instants donnés. Il sera donc possible de prédire le comportement du système dans un futur correspondant à plusieurs pas.

## 5.1 ESTIMATEUR NEURONAL

### 5.1.1 Modèle de l'avion Boeing 747

Le système dont il faut estimer les variables d'état est l'avion Boeing 747 en commande du mouvement longitudinal (Bryson, 1994); stable en boucle ouverte. Le

vecteur d'état  $x_k$  est donné par:

$$x_k = \begin{bmatrix} u_a \\ w_a \\ q_a \\ \theta_a \end{bmatrix} \quad (5.1)$$

où

$u_a$  : vitesse sur l'axe de roulis

$w_a$  : vitesse sur l'axe de lacet

$q_a$  : vitesse de tangage

$\theta_a$  : angle de tangage

Son modèle dans l'espace d'état est exprimé par le système d'équations (5.2).

$$\begin{cases} x_{k+1} = Fx_k + Bu_k + G\omega_k \\ y_k = Cx_k + v_k \end{cases} \quad (5.2)$$

où les matrices  $F$ ,  $B$ ,  $G$  et  $C$  sont données par:

$$F = \begin{bmatrix} 0.9954 & 0.0035 & -0.0011 & -0.0321 \\ -0.0318 & 0.7918 & 0.1373 & 0.0005 \\ 0.0221 & -0.3087 & 0.7170 & -0.0004 \\ 0.0011 & -0.0154 & 0.0858 & 1.0000 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.0057 & 0.9977 \\ -1.0075 & -0.0159 \\ -9.3999 & 0.0110 \\ -0.4700 & 0.0006 \end{bmatrix}$$

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

### 5.1.2 Structure et caractéristiques du réseau de neurones utilisé

Le réseau, par lequel nous avons remplacé le gain du filtre de Kalman, est celui qui apparaît à la figure 5.1. La loi d'adaptation des poids est toujours celle donnée ci-dessous:

$$w_{ij}(k+1) = w_{ij}(k) + \eta e^t(k) CF[F \frac{\partial \hat{x}_{k-2}}{\partial w_{ij}} + \frac{\partial \Gamma_{k-1}}{\partial w_{ij}}] \quad (5.3)$$

L'adaptation des poids se fait à chaque itération.

Le réseau utilisé qui apparaît à la figure 5.1 a deux entrées (soit  $n_a = 2$ ). Il devra donc avoir cinq neurones à sa couche cachée. Dans cette analyse, nous avons utilisé dix neurones. Le nombre d'états de l'avion étant de 4 (quatre), il a fallu quatre sorties au réseau de neurones ce qui fixe à 4 (quatre) le nombre de neurones de la couche de sortie.

Dans ce cas particulier, le calcul de la variation de la sortie du réseau par rapport au vecteur des poids ( $\frac{\partial \Gamma_{k-1}}{\partial w}$ ) se réduit à:

$$\Gamma_i(k) = \sum_{j=1}^{10} z_{ij} f_j \left( \sum_{j_1=1}^2 [v_{j_1 j} E_e(j_1)] + b_j \right) \quad (5.4)$$

$\Gamma(k) = [\Gamma_1(k), \Gamma_2(k), \Gamma_3(k), \Gamma_4(k)]^t$  (où  $E_e$  est un vecteur dont les composantes sont les deux entrées du réseau)

ce qui permet de déduire:

$$\frac{\partial \Gamma_i(k)}{\partial z_{ij}} = f_j \left( \sum_{j_1=1}^2 [v_{j_1 j} E_e(j_1)] + b_j \right) \quad (5.5)$$

et

$$\frac{\partial \Gamma_i(k)}{\partial v_{j_1 j}} = z_{ij} \dot{f}_j \left( \sum_{j_3=1}^2 [v_{j_3 j} E_e(j_3)] + b_j \right) E_e(j_1) \quad (5.6)$$

où  $v_{ij}$  et  $z_{ij}$  sont respectivement les poids de la couches cachée et ceux de la couche de sortie.

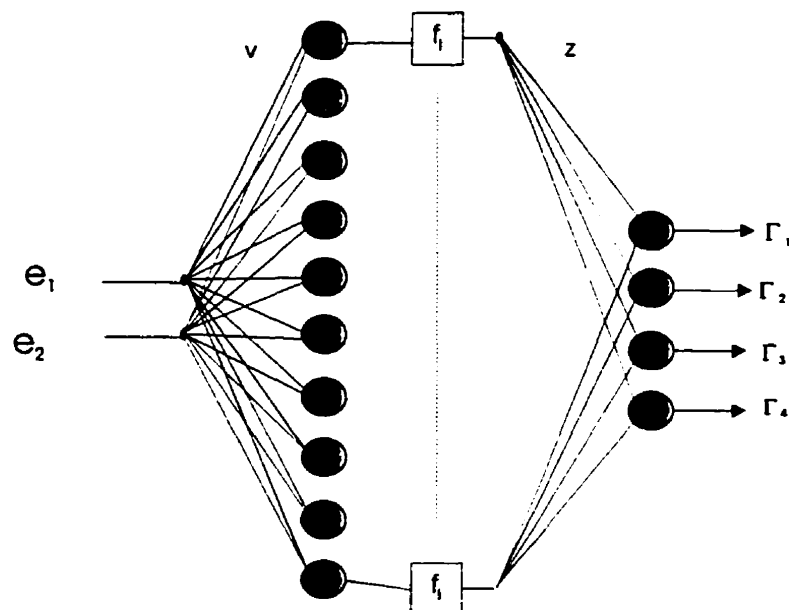


Figure 5.1: Structure du Réseau estimateur utilisé

Tableau 5.1: Caractéristiques du réseau estimateur utilisé

	couche cachée	couche de sortie
nombre de neurones	10	4
fonction d'activation	sigmoïde	linéaire

### 5.1.3 Données de simulation

#### 5.1.3.1 Les bruits sont blancs et de covariances connues

Les bruits  $\omega_k$  et  $v_k$  sont blancs de moyennes nulles et de variances respectives 9 et 1. Pour exciter davantage le système, on choisit les signaux d'entrée  $u_1(k)$  (une onde carrée d'amplitude 0.01 rad) et  $u_2(k)$  (un signal sinusoïdal d'amplitude 0.005 rad) (figure 5.2). Les neurones de la couche de sortie et les biais ont un coefficient d'apprentissage  $\eta$  égal à 0.01 alors que celui de la couche cachée est de 0.04. Le choix de ces valeurs est justifié par le fait que les données d'apprentissage étant non normalisées, il est souvent préférable d'utiliser des coefficients d'apprentissage



faibles. Les résultats de simulation sont au chapitre suivant.

### 5.1.3.2 Les bruits sont non blancs et non gaussiens de covariances non connues

Dans cette analyse, les bruits non blancs et non gaussiens sont pris en compte et l'intérêt est porté sur le comportement des deux estimateurs. Le filtre neuronal ne requière pas la connaissance des bruits pendant l'estimation qui n'est basée que sur la minimisation de l'erreur  $y_k - \hat{y}_k$  par action sur les poids. Pour matérialiser l'hypothèse des statistiques inconnues, dans le souci d'implanter le filtre de Kalman, on admet que le filtre de Kalman ne peut fonctionner qu'en prédicteur. Son équation se réduit à celle ci-dessous:

$$x_{k+1/k} = F^k x_{0/0} + \sum_{i=1}^k F^{k-i} B u_{i-1} \quad (5.7)$$

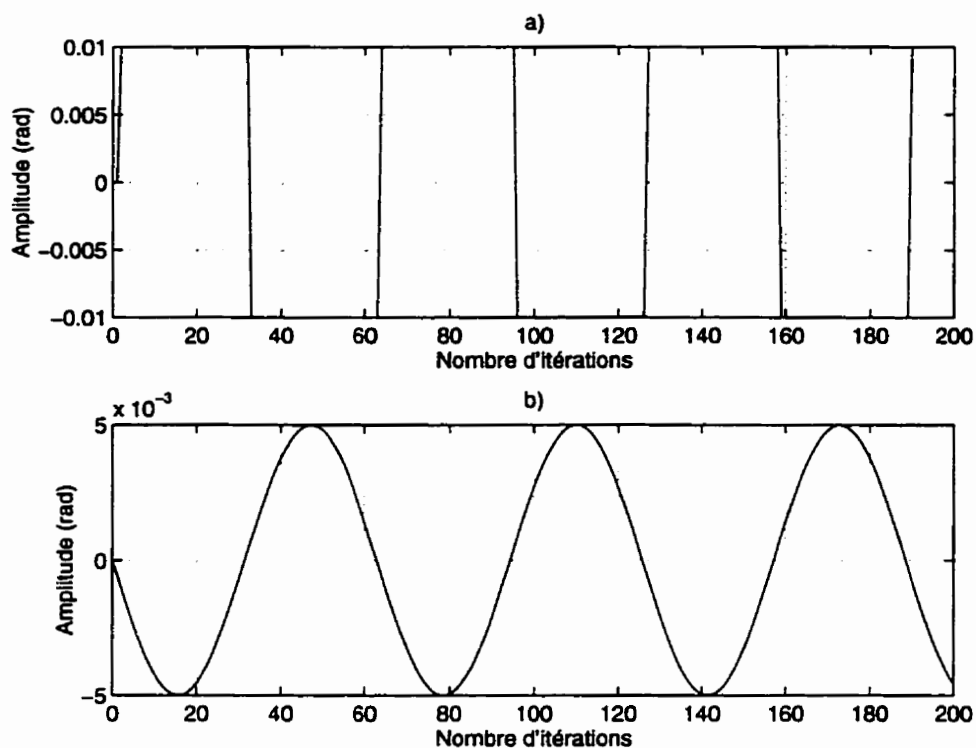
où  $x_{0/0}$  est la condition initiale

Des bruits de modèle et de mesure non blancs et non gaussiens ont été générés, dans chacun des cas, en prenant le sinus d'un bruit blanc et en le multipliant ensuite par ce bruit puis en considérant à chaque instant, l'exponentiel du produit soit:

$$\bar{\omega}_k = \exp(\omega_k \sin(\omega_k)) \quad (5.8)$$

$$\bar{v}_k = \exp(v_k \sin(v_k)) \quad (5.9)$$

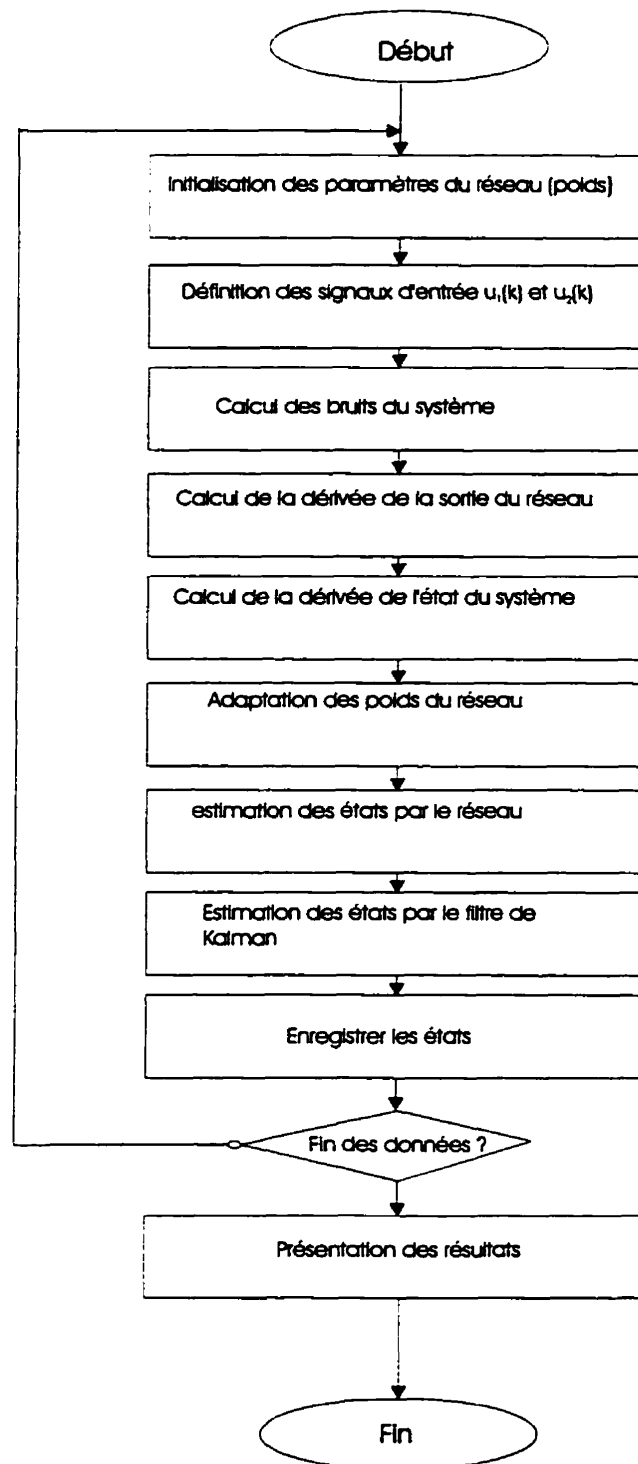
Dans ce cas, il serait difficile d'estimer les statistiques de ces bruits dans un système réel.



**Figure 5.2:** Signaux d'entrée utilisés pour l'estimation des états du Boeing 747; a) Entrée  $u_1$  (rad); b) Entrée  $u_2$  (rad).

#### 5.1.4 Structure du programme d'estimation

L'organigramme de programmation est celui décrit à la figure 5.3. On utilise un réseau dont les caractéristiques et la structure apparaissent au tableau 5.1 et à la figure 5.1.



**Figure 5.3:** Organigramme de programmation de l'estimateur neuronal et de comparaison avec le filtre de Kalman

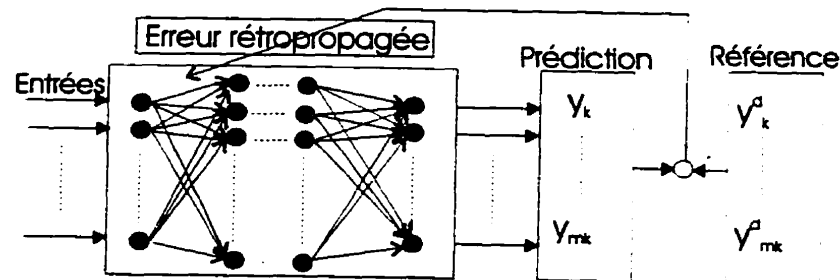


Figure 5.4: Réseau prédicteur en phase d'apprentissage.

## 5.2 PRÉDICTEUR NEURONAL

Le principe de prédiction est basé sur la capacité des réseaux de neurones artificiels de pouvoir procéder à l'apprentissage de la dynamique des systèmes par ajustage de leurs poids en fonction des données d'entrée et de sortie. Un réseau de neurones que ce soit en phase d'apprentissage (les poids varient figure 5.4) ou en phase de rappel (les poids restent constants figure 5.5) fonctionne en prédiction; pour cette raison, tout réseau de neurones est un prédicteur.

### 5.2.1 Le système en étude

Le système considéré est formé de trois réservoirs en cascade de stockage d'eau de l'Hydro-Québec (Baskatong, Maniwaki et Cabonga). L'objectif visé est la prédiction des apports naturels aux réservoirs.

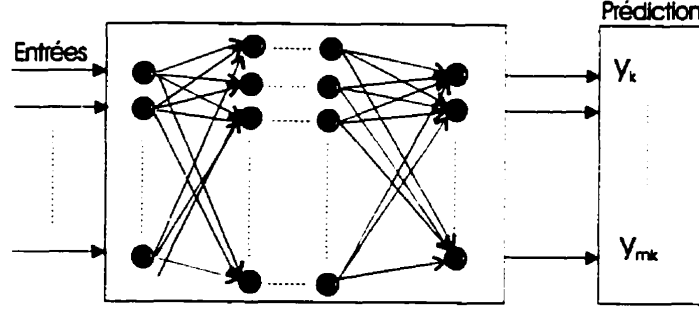


Figure 5.5: Réseau prédicteur en rappel.

#### 5.2.1.1 Modèle des réservoirs

Les prédictions, à chaque instant, sont régies par les équations ci-dessous (BERGMAN et DELLEUR, 1985):

$$y_1(k) = \sum_{j=1}^{n_r} a_j(k)y_1(k-j) + a_{n_r+j}(k)y_2(k-j) + a_{2n_r+j}(k)y_3(k-j) + e_{1k} \quad (5.10)$$

$$y_2(k) = \sum_{j=1}^{n_r} b_j(k)y_1(k-j) + b_{n_r+j}(k)y_2(k-j) + b_{2n_r+j}(k)y_3(k-j) + e_{2k} \quad (5.11)$$

$$y_3(k) = \sum_{j=1}^{n_r} c_j(k)y_1(k-j) + c_{n_r+j}(k)y_2(k-j) + c_{2n_r+j}(k)y_3(k-j) + e_{3k} \quad (5.12)$$

où  $n_r$  est le nombre de mesures passées et  $y_1(k)$ ,  $y_2(k)$ ,  $y_3(k)$ , sont les mesures des apports naturels.

Le système formé par les équations ci-dessus peut se mettre sous la forme compacte suivante:

$$y_k = \phi_k \theta_k + e_k \quad (5.13)$$

avec  $\phi_k$  la matrice d'informations entrée/sortie et  $\theta_k$  le vecteur des paramètres du modèle. N'ayant de données que les niveaux des réservoirs en fonction du temps, il fallait d'abord déterminer les paramètres  $a_i$ ,  $b_i$  et  $c_i$ . Pour ce faire, il fallait trouver

combien de valeurs passées ( $n_r$ ) le modèle devait comporter. Dans les paragraphes qui suivent, une justification du choix retenu est faite.

**Modèle basé sur deux mesures passées ( $n_r = 2$ )** Les équations (5.10), (5.11), (5.12) peuvent se mettre sous la forme:

$$\begin{cases} x_{k+1} = Fx_k + w_k \\ y_k = Cx_{k+1} + v_k \end{cases} \quad (5.14)$$

où

$$\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \end{bmatrix} = \begin{bmatrix} y_1(k-2) \\ y_1(k-1) \\ y_2(k-2) \\ y_2(k-1) \\ y_3(k-2) \\ y_3(k-1) \end{bmatrix} \quad (5.15)$$

et les matrices  $F$  et  $C$  sont:

$$F = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ a_2 & a_1 & a_4 & a_3 & a_6 & a_5 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ b_2 & b_1 & b_4 & b_3 & b_6 & b_5 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ c_2 & c_1 & c_4 & c_3 & c_6 & c_5 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

**Tableau 5.2:** Paramètres du modèle des réservoirs (modèle basé sur deux mesures passées) identifiés par le filtre de Kalman étendu

$a_1$	$b_1$	$c_1$
1.9979	-0.0135	0.1233
$a_2$	$b_2$	$c_2$
-1.0475	0.0010	-0.1274
$a_3$	$b_3$	$c_3$
-0.2403	1.9581	-0.0214
$a_4$	$b_4$	$c_4$
0.4652	-0.9798	0.0414
$a_5$	$b_5$	$c_5$
-2.8309	0.1518	0.3638
$a_6$	$b_6$	$c_6$
2.8015	0.0033	0.6311

**Modèle basé sur trois mesures passées ( $n_r = 3$ )** Nous considérons toujours le système (5.14); mais les états deviennent:

$$\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \\ x_5(k) \\ x_6(k) \\ x_7(k) \\ x_8(k) \\ x_9(k) \end{bmatrix} = \begin{bmatrix} y_1(k-3) \\ y_1(k-2) \\ y_1(k-1) \\ y_2(k-3) \\ y_2(k-2) \\ y_2(k-1) \\ y_3(k-3) \\ y_3(k-2) \\ y_3(k-1) \end{bmatrix} \quad (5.16)$$

et les matrices  $F$  et  $C$  sont dans ce cas données comme ci-dessous:

$$F = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_6 & a_5 & a_4 & a_9 & a_8 & a_7 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ b_3 & b_2 & b_1 & b_6 & b_5 & b_4 & b_9 & b_8 & b_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ c_3 & c_2 & c_1 & c_6 & c_5 & c_4 & c_9 & c_8 & c_7 \end{bmatrix}$$

**Tableau 5.3:** Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu.

$a_1$	$b_1$	$c_1$
3.3364	-0.1270	0.2656
$a_2$	$b_2$	$c_2$
-1.7718	0.0716	-0.1827
$a_3$	$b_3$	$c_3$
-0.5511	0.0418	-0.0817
$a_4$	$b_4$	$c_4$
0.7777	2.1825	0.0686
$a_5$	$b_5$	$c_5$
-0.5200	-1.8398	-0.0454
$a_6$	$b_6$	$c_6$
-0.0372	0.6587	-0.0036
$a_7$	$b_7$	$c_7$
-11.3345	1.6689	-0.6580
$a_8$	$b_8$	$c_8$
-3.6121	-1.1711	-0.0401
$a_9$	$b_9$	$c_9$
14.4419	-0.3628	1.6532

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

#### Choix du nombre de mesures passées ( $n_r$ )

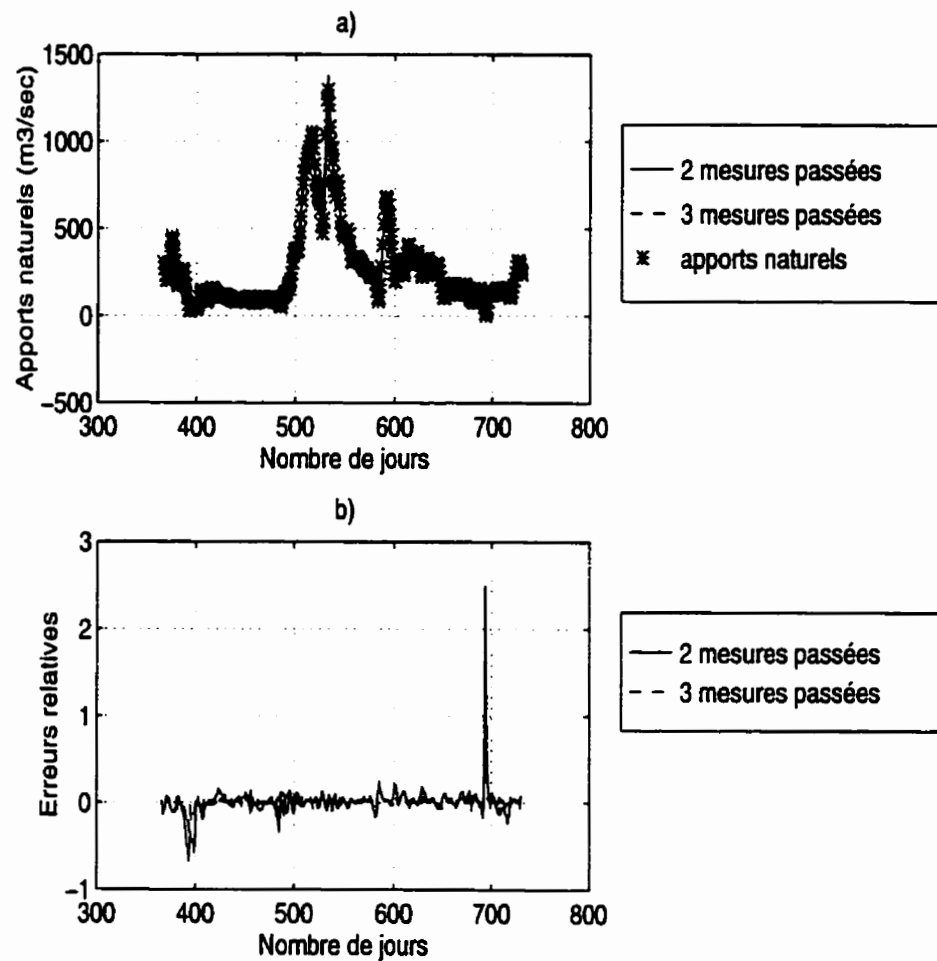
A partir des figures 5.6;5.7 et 5.8; on peut remarquer que les modèles des réservoirs basés sur trois ou sur deux mesures passées se comportent de la même façon et représentent au mieux le système réel. Il faut cependant noter que le modèle sur trois mesures passées a été utilisé parce qu'il s'approche davantage du comportement réel à cause de la faiblesse de son erreur relative  $E_r$ , donnée ci-dessous:

$$E_r(k) = \frac{y(k) - \hat{y}(k)}{y(k)} \quad (5.17)$$

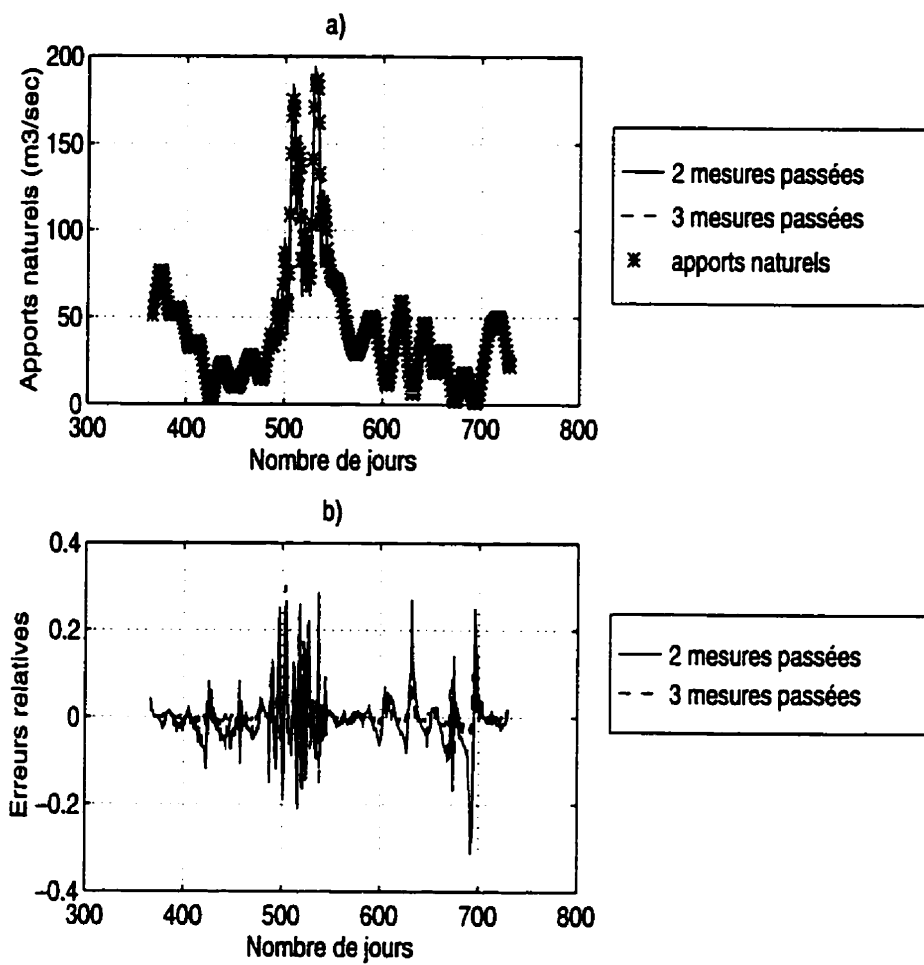
Il existe certainement plusieurs autres types de calculs d'erreurs normalisées pour les fins de comparaison, mais celle que nous adoptons dans cette étude est l'erreur



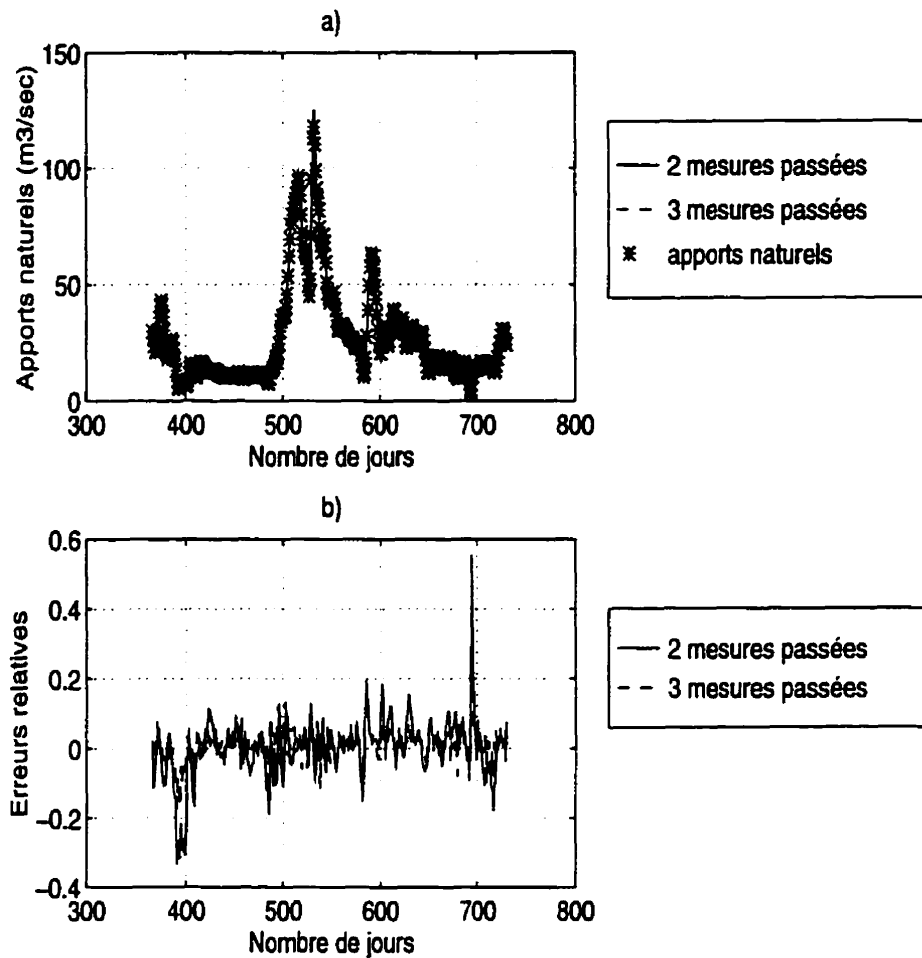
relative.



**Figure 5.6:** a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Baskatong)



**Figure 5.7:** a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Cabonga)



**Figure 5.8:** a) Apports naturels réels et identifiés par le filtre de Kalman étendu à partir des modèles à 3 et 2 mesures passées; b) Erreurs relatives (Maniwaki)

### 5.2.2 Données de Simulation

Nous disposons des données donnant les apports naturels (échantillonnage par jour) aux réservoirs sur une période de 39 ans. Ces données sont bien évidemment bruitées; le type et les statistiques des bruits restent toutefois inconnus. Dans

Tableau 5.4: Paramètres d'apprentissage

dis	maxepoch	erreur	mingr	mu	muinc	mudec	mumax
10	100	0.3	0.0001	0.001	10	0.1	$10^{10}$

le cadre de cette simulation, nous nous intéressons aux trois premières années. Les données pour les deux premières années (730 par réservoir) servent de données d'apprentissage tandis que celles de la troisième année nous permettront de vérifier la généralisation du filtre neuronal (rappel). Ces données sont celles apparaissant aux figures 5.11; 5.9; 5.10. L'apprentissage avait été réalisé sur UNIX SPARCstation 5 (Wiener). Pour couvrir 1200 "epoch", la durée moyenne d'apprentissage était de huit heures (en fonctionnement continu). Il faut rappeler ici que malgré ce long temps d'apprentissage, une parfaite convergence des poids n'avait pas été atteinte. L'erreur d'apprentissage continuait de décroître. Ce manque convergence rapide des poids était due à la non normalisation des données d'apprentissage. En outre, l'apprentissage se faisait par "batch" de  $730 \times 3$ . Il ne fallait pas non plus normaliser les données au risque d'introduire d'autres types de bruits à ceux que contenaient déjà les données. Il était nécessaire d'opter pour la limitation du nombre d'"epoch" pendant l'apprentissage. Le tableau 5.4 montre les paramètres d'apprentissage où:

- dis : fréquence d'affichage des paramètres
- maxepoch : nombre maximal des "epoch"
- erreur : valeur de l'erreur désirée
- mingr : valeur minimale admissible du gradient
- mu : valeur initiale du moment

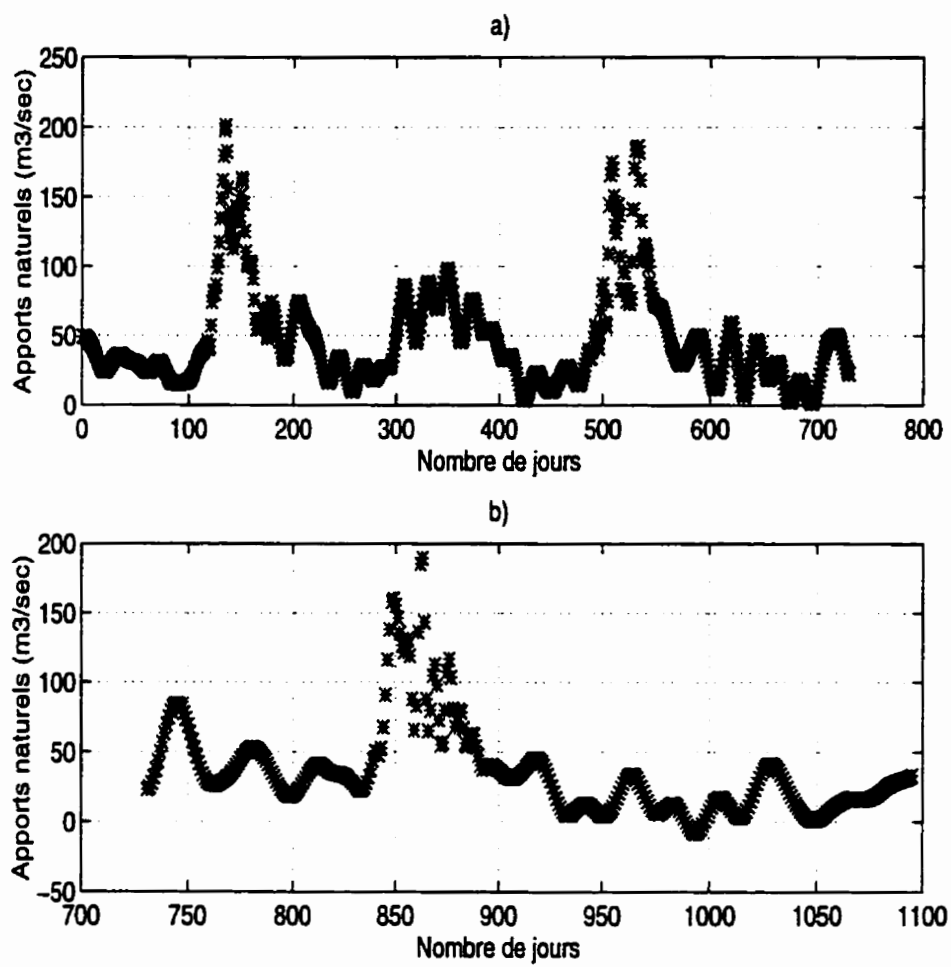
- muinc : pas d'incrémentation du moment
- mudec : pas de décrémentation du moment
- mumax : valeur maximale du moment.

Que ce soit pendant la phase d'apprentissage ou de rappel, il s'agira de prédire l'apport naturel du réservoir à un jour donné à partir des apports naturels des trois jours précédents (prédire  $y(k)$  à partir de  $y(k-1)$ ,  $y(k-2)$  et  $y(k-3)$ ). Pour les trois réservoirs, à un instant donné  $k$ , il y aura douze entrées donc neuf mesures passées des apports naturels puis trois erreurs provenant de la différence  $y_k - \hat{y}_k$  et trois sorties.

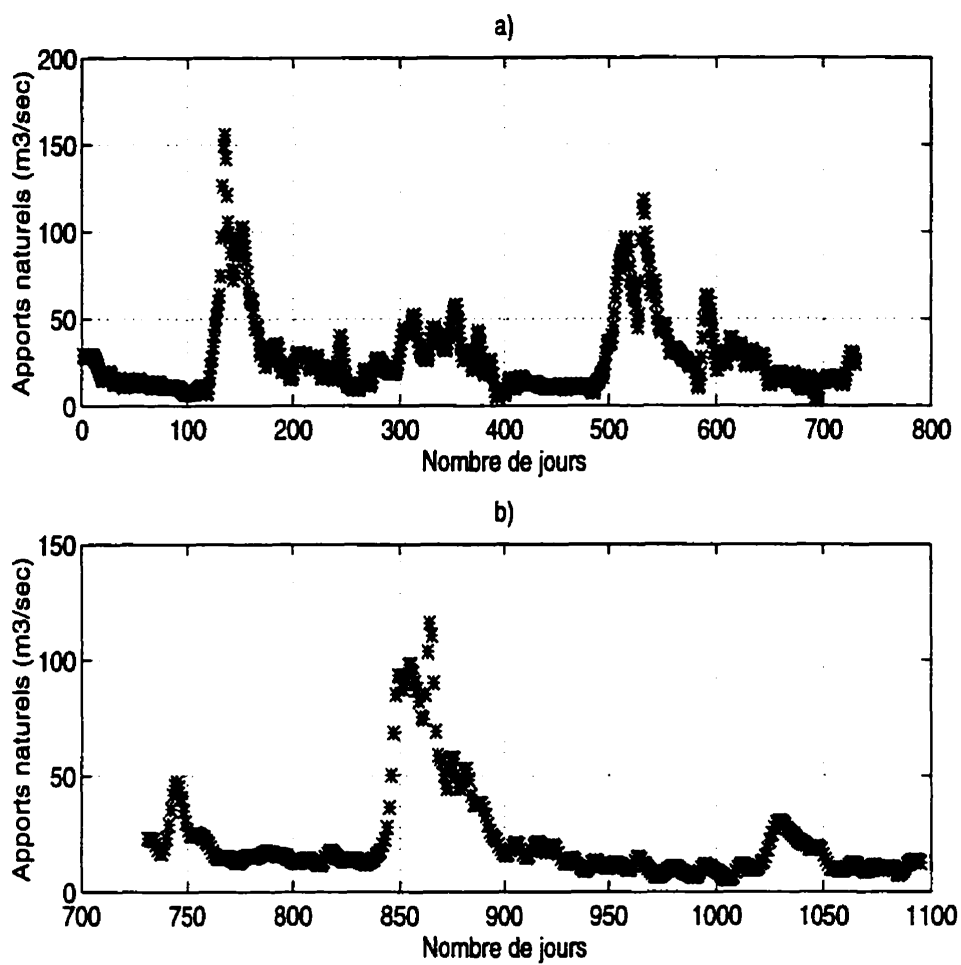
Le vecteur  $w_k$  est une séquence de bruit que nous supposons blanc, gaussien, de moyenne nulle et de covariance  $Q_k$ . Par ailleurs,  $v_k$  est un vecteur nul; puisque les  $y_k$  sont des données déjà bruitées. Pour déterminer  $Q_k$  et  $R_k$  (covariance de l'erreur d'estimation) nous avons utilisé la méthode de Assaf et Quick, (1991) qui proposent de calculer  $R_k$  et la matrice  $Q_k(i, j)$  comme ci-dessous:

$$R_k = (0.15y_k)^2 \quad (5.18)$$

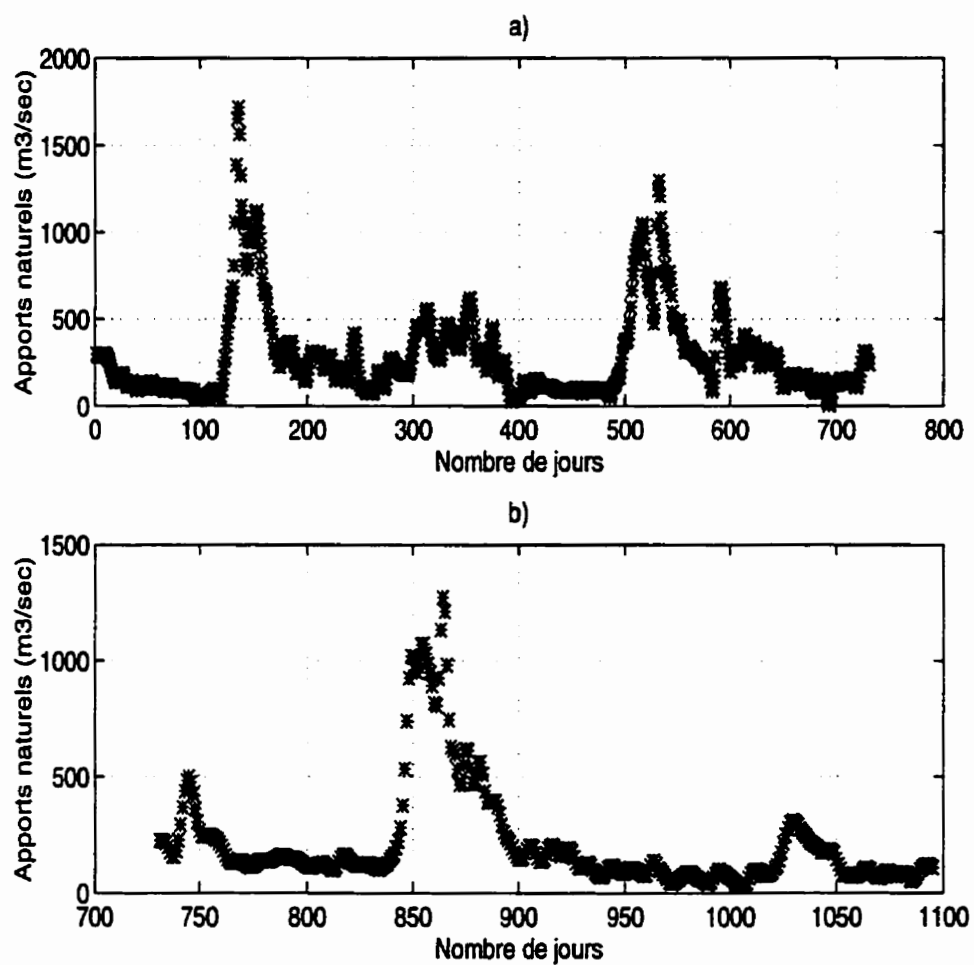
$$Q_k(i, j) = (0.1F_{ij})^2 \delta_{ij} \quad (5.19)$$



**Figure 5.9:** Données sur deux ans du réservoir Cabonga utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel



**Figure 5.10:** Données sur deux ans du réservoir Maniwaki utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel



**Figure 5.11:** a) Données sur deux ans du réservoir Baskatong utilisées lors de l'apprentissage; b) Données sur un an utilisées lors du rappel



### 5.2.3 Structure et caractéristiques du réseau de neurones utilisé

Le réseau prédicteur utilisé dans le cadre de cette simulation est celui dont la structure et les caractéristiques apparaissent respectivement à la figure 5.12 et au tableau 5.5. Il comporte une seule couche cachée de vingt-cinq neurones et une couche de sortie de trois neurones identifiant les apports naturels des trois réservoirs. Le réseau comporte douze entrées formées des valeurs aux trois instants précédents des apports naturels de chaque réservoir et du vecteur d'erreur  $y_k - \hat{y}_k$ . La fonction  $f_j$  est linéaire alors que  $f_s$  est non linéaire de type sigmoïde et donnée par la relation:

$$f_s(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (5.20)$$

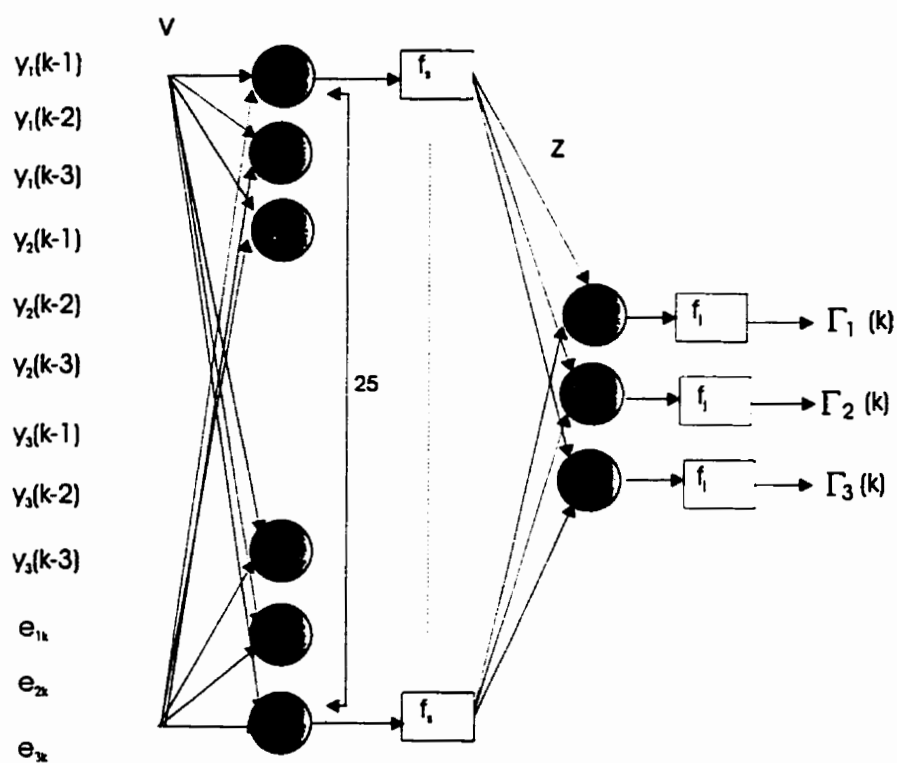
Les  $v_{ij}$  sont les poids de la couche cachée et les  $z_{js}$  ceux de la couche de sortie.

### 5.2.4 Structure du programme de prédiction

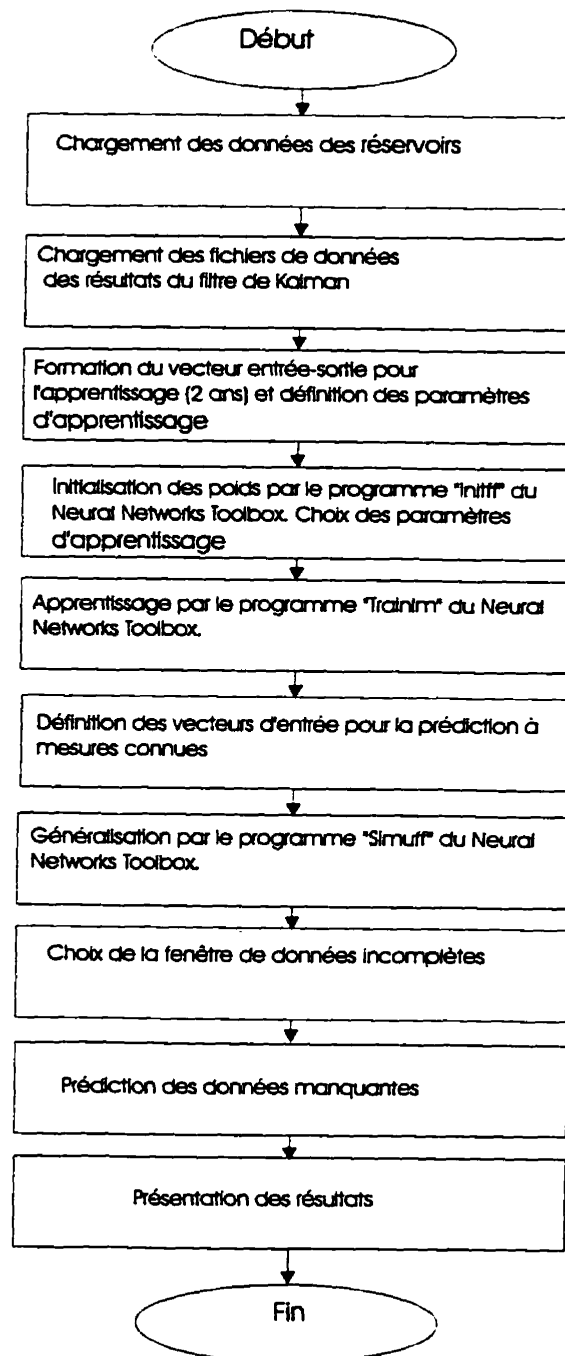
La méthode d'apprentissage utilisée est la rétropropagation basée sur l'approximation de Levenberg Marquardt. Cette méthode est programmée dans le "Neural Networks Toolbox", (Demuth et Mark Beale, 1994) de Matlab mais a été modifiée pour qu'elle puisse s'adapter à la présente simulation. L'organigramme autour duquel est bâti le programme de programmation apparaît à la figure 5.13.

**Tableau 5.5:** Caractéristiques du réseau prédicteur utilisé

	couche d'entrée	couche de sortie
nombre de neurones	25	3
fonction d'activation	sigmoïde	linéaire



**Figure 5.12:** Réseau prédicteur utilisé



**Figure 5.13:** Organigramme de programmation du prédicteur neuronal et de comparaison avec le filtre de Kalman

## Conclusion

Le présent chapitre a permis de voir que le filtre neuronal peut jouer le rôle d'un estimateur ou d'un prédicteur. Il est possible d'estimer les états de l'avion Boeing 747 d'une part et de prédire les apports naturels des trois réservoirs de l'Hydro-Québec d'autre part. Aussi est-il nécessaire de comparer ses performances à celles du filtre de Kalman. Les chapitres 6 et 7 suivants, qui donnent les résultats de simulation, établissent cette comparaison et montrent l'importance du filtre neuronal. Pour mieux apprécier le comportement du filtre neuronal, nous nous proposons d'analyser les conditions de fonctionnement des deux filtres lorsque le niveau de bruit dans les données est élevé.

Dans le cas de l'estimateur, il s'agira de voir l'effet des bruits non blancs sur les filtres. Le filtre de Kalman dans le contexte des statistiques inconnues, mathématiquement ne peut plus donner de bons résultats. L'estimateur neuronal qui ne minimise que l'erreur pourra produire de résultats satisfaisants. L'intérêt est de voir si dans le cas d'un rapport bruit/signal élevé le comportement du filtre neuronal est toujours admissible. Par ailleurs, si la matrice d'état  $F$  est perturbée les deux filtres restent-ils robustes? Aussi nous augmenterons le degré de perturbation de la matrice  $F$  et observerons l'effet sur la performance de l'estimateur neuronal. En outre, dans le cas de la prédiction, nous verrons jusqu'à quel pourcentage, bruit/signal, les deux prédicteurs peuvent donner de bons résultats.

## Chapitre 6

# ANALYSE COMPARATIVE DU FILTRE DE KALMAN ET DE L'ESTIMATEUR NEURONAL

### Introduction

Ce chapitre est une plate-forme de comparaison entre le filtre de Kalman et l'estimateur neuronal. A partir des résultats de simulation des deux estimateurs nous dégagerons certaines conclusions. Tout au long de la comparaison, le filtre de Kalman est pris pour référence dans les conditions où son utilisation est possible. En effet, le filtre de Kalman est optimal si les bruits sont blancs et de statistiques connues. Par ailleurs, les critères qui définissent sa stabilité et sa convergence sont clairement connus. Aussi les résultats donnés par l'estimateur neuronal seront analysés par rapport au comportement du filtre de Kalman. Ainsi, dans chaque analyse, il faudra prêter attention à l'évolution temporelle des erreurs moyennes des deux estimateurs.

On supposera d'une part, que les statistiques des bruits sont connues et d'autre part, qu'elles sont inconnues. Pour mener à profit cette analyse, on s'intéressera au

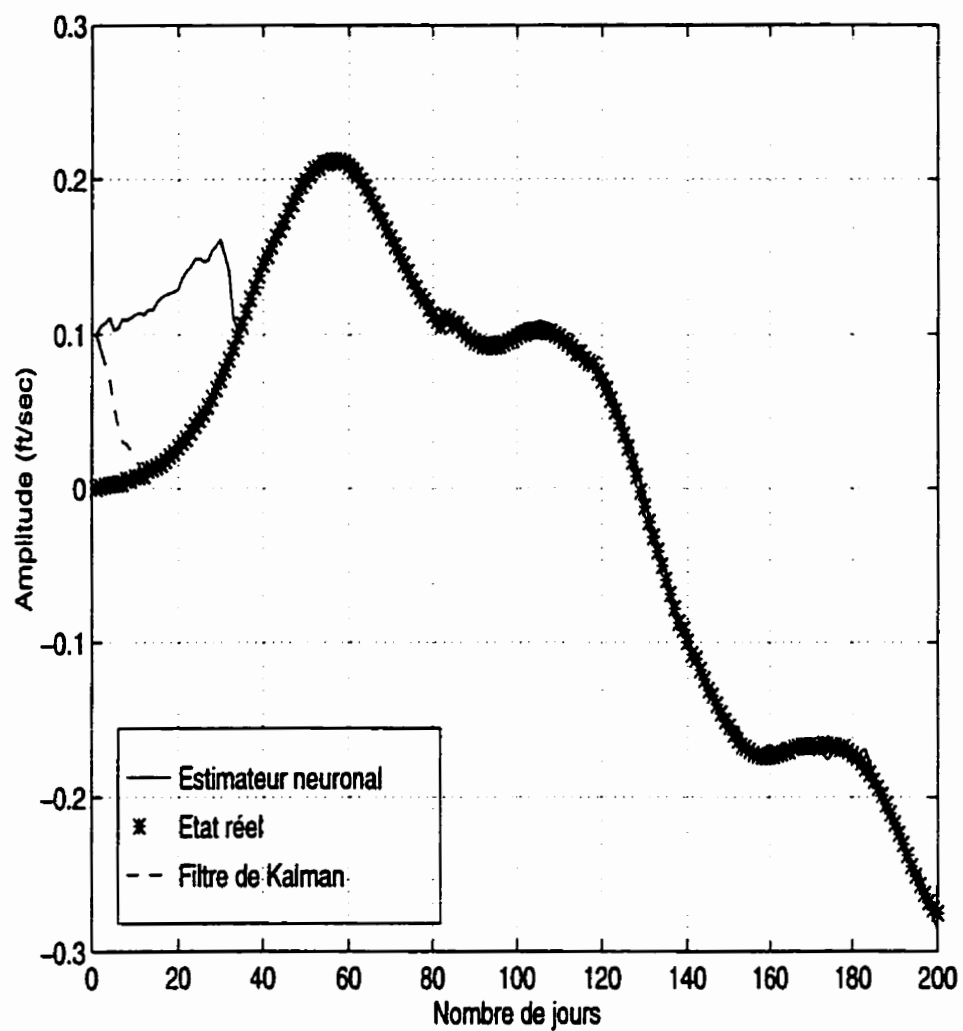
cas où les bruits qui affectent le système (avion) sont blancs et gaussiens. En outre, il s'agira de voir le comportement des estimateurs si les bruits sont non blancs et non gaussiens de valeur efficace unitaire. Le cas qui retiendra aussi l'attention est celui où le rapport bruit/signal est non négligeable. L'on observera le comportement de l'estimateur neuronal lorsque le rapport bruit/signal est de 5% et 10%. Par ailleurs, il sera question de voir si les deux estimateurs sont robustes. On fera varier les modules des pôles du système en boucle ouverte de 5% et de 10%.

## **6.1 ANALYSE COMPARATIVE DES DEUX ESTIMATEURS**

### **6.1.1 Les bruits sont blancs et gaussiens de statistiques connues**

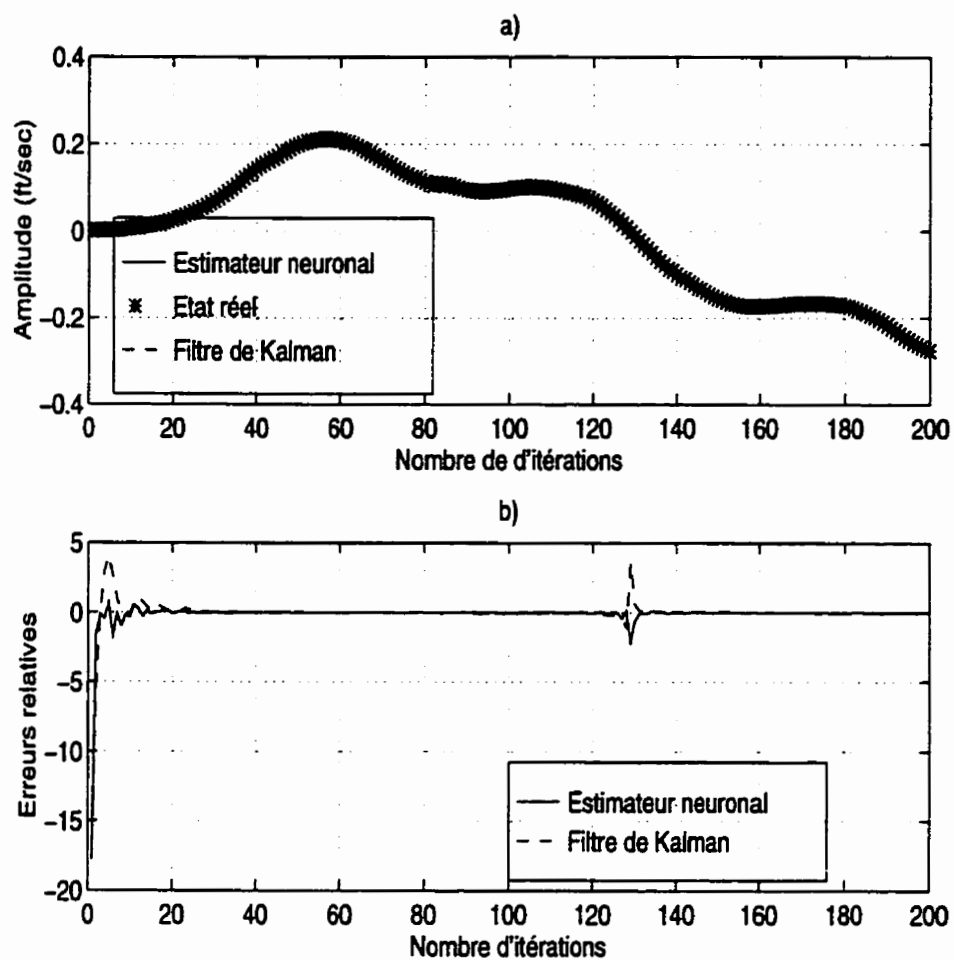
Les figures 6.2, 6.3, 6.4, 6.5 des pages suivantes, montrent les estimés des variables d'état du mouvement longitudinal du Boeing 747 produits par le filtre de Kalman et par l'estimateur neuronal sous l'influence des entrées  $u_1(t)$  et  $u_2(t)$ . L'estimateur neuronal a un comportement quasi identique à celui du filtre de Kalman. Les erreurs relatives sont globalement admissibles ce qui permet de dire que les états réels sont proches de ceux estimés par les deux filtres. Ce qu'on peut en outre remarquer, c'est la rapide convergence du filtre de Kalman par rapport au filtre neuronal. La figure 6.1 montre les estimés des deux filtres de l'état  $x_1$  pour une condition initiales égale à 0.1. Le filtre de Kalman converge après quelques itérations alors que le filtre neuronal prend plus de temps à converger. Dans la suite de l'étude, nous poserons les conditions initiales, aussi bien du filtre neuronal que du filtre de Kalman à 0.001

pour tous les états. Ce choix permet d'avoir une convergence rapide des deux filtres.

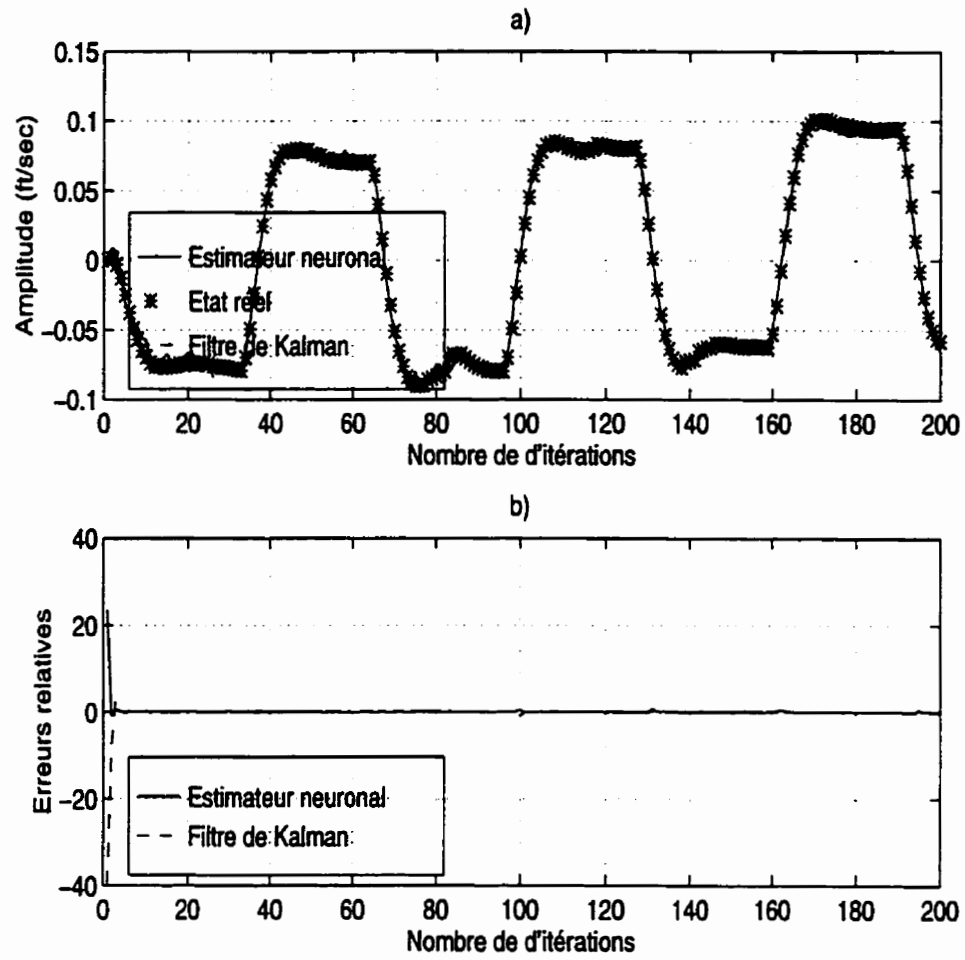


**Figure 6.1:** Estimés de l'état  $x_1$  (vitesse sur l'axe de roulis) par le filtre neuronal et le filtre de Kalman pour une condition initiale égale à 0.1 ft/sec.

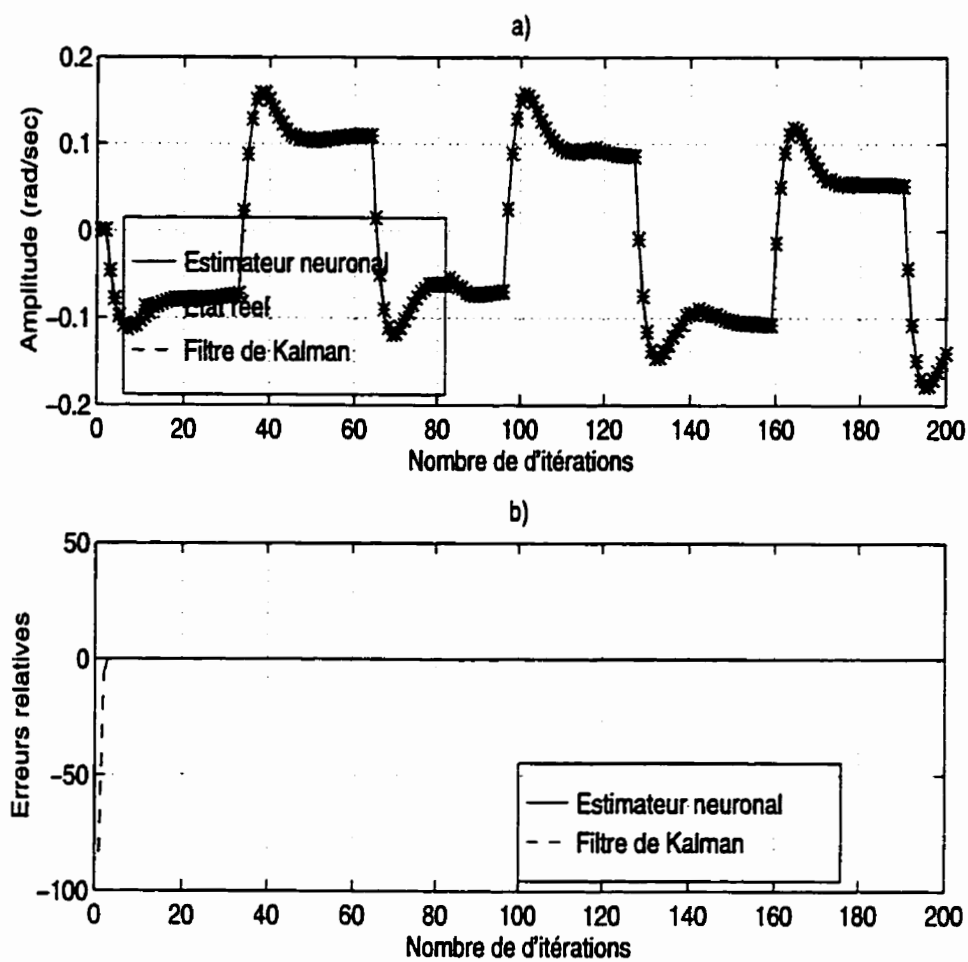




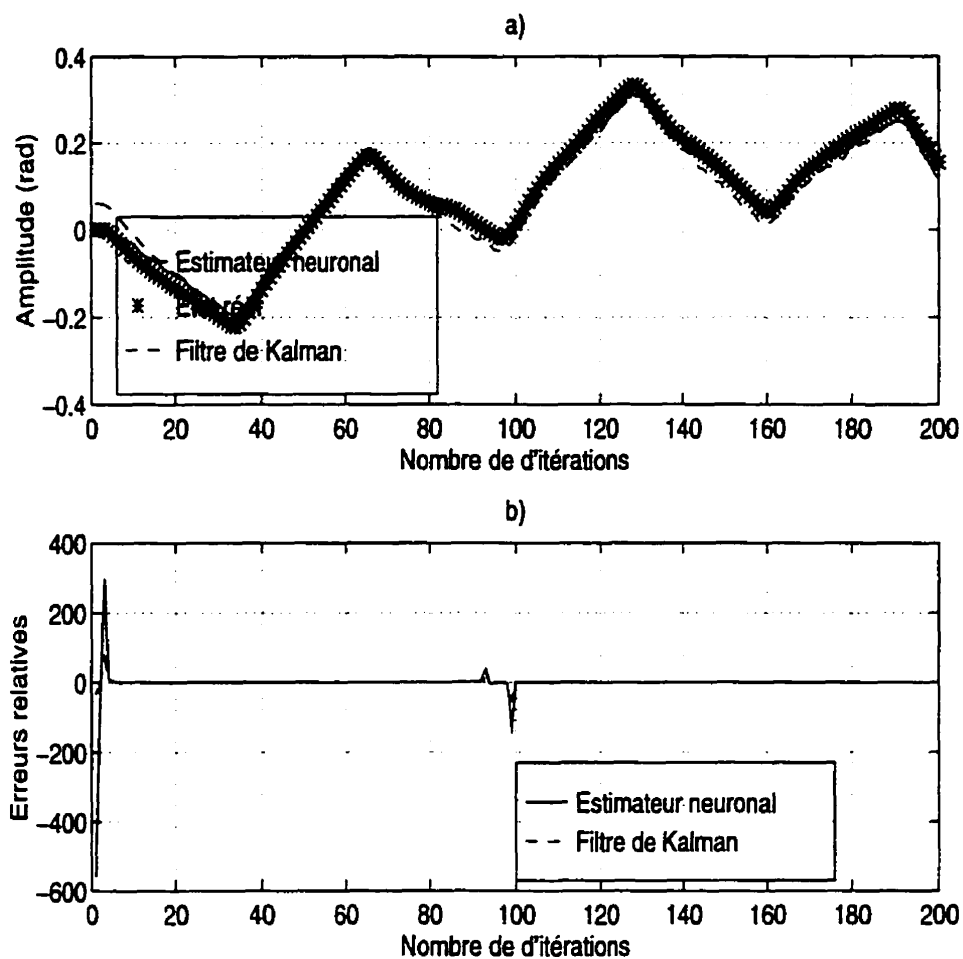
**Figure 6.2:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives.



**Figure 6.3:** a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives.



**Figure 6.4:** a) Etat  $x_3$  (vitesse de tangage) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives.



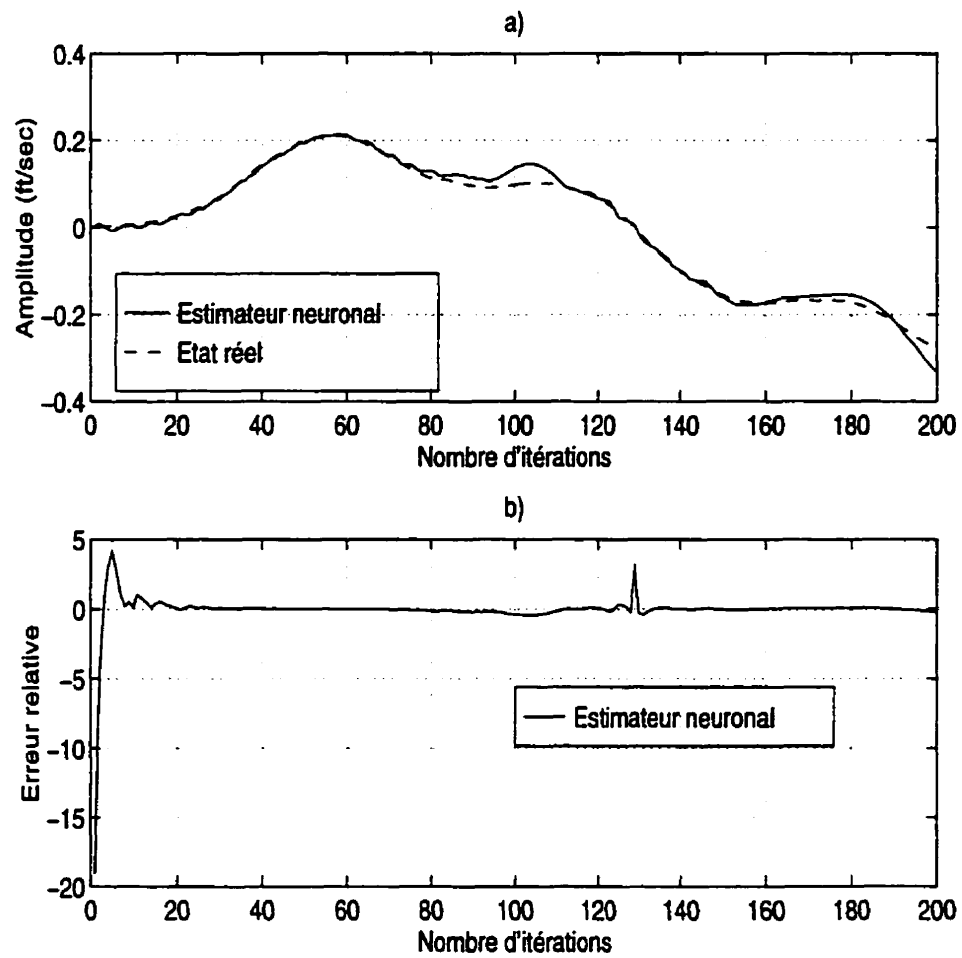
**Figure 6.5:** a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal et par le filtre de Kalman; b) Erreurs relatives.

## **6.1.2 Les bruits sont non blancs et non gaussiens de statistiques inconnues**

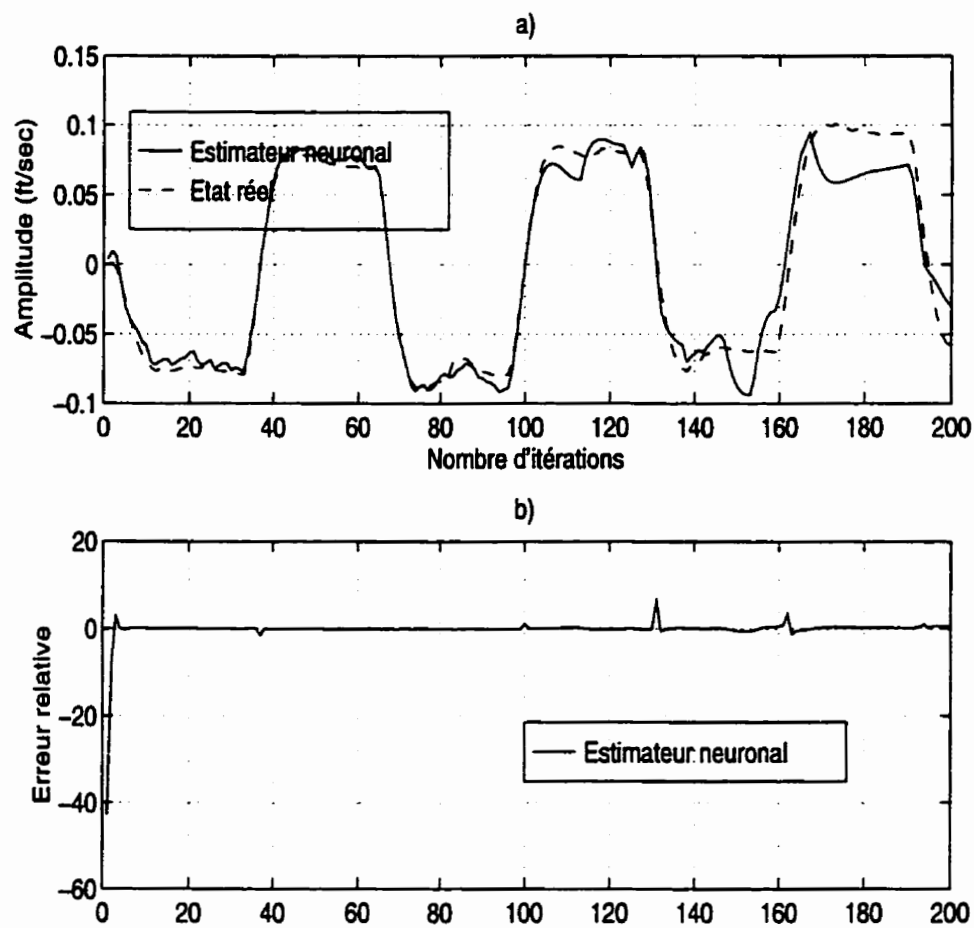
### **6.1.2.1 L'intensité des bruits est faible**

Les figures 6.6, 6.7, 6.8 et 6.9 des pages suivantes mettent en évidence les états de l'avion estimés par le réseau. On peut voir concrètement que le réseau suit assez bien la dynamique des états bien que les bruits soient non blancs et non gaussiens. Les statistiques des bruits étant inconnues, le filtre de Kalman ne peut plus être utilisé.

L'évolution temporelle des estimés des états de l'avion permet de constater que l'estimateur neuronal présente un grand avantage, par rapport au filtre de Kalman, pour l'estimation des états des systèmes dynamiques dans un contexte de bruits de modèle et de mesure non blancs et surtout de statistiques inconnues. Ce comportement de l'estimateur neuronal peut laisser croire qu'il donnerait toujours de bons résultats; aussi allons-nous voir sa réponse si le rapport bruit/signal est élevé aux paragraphes suivants.



**Figure 6.6:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative.



**Figure 6.7:** a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative .

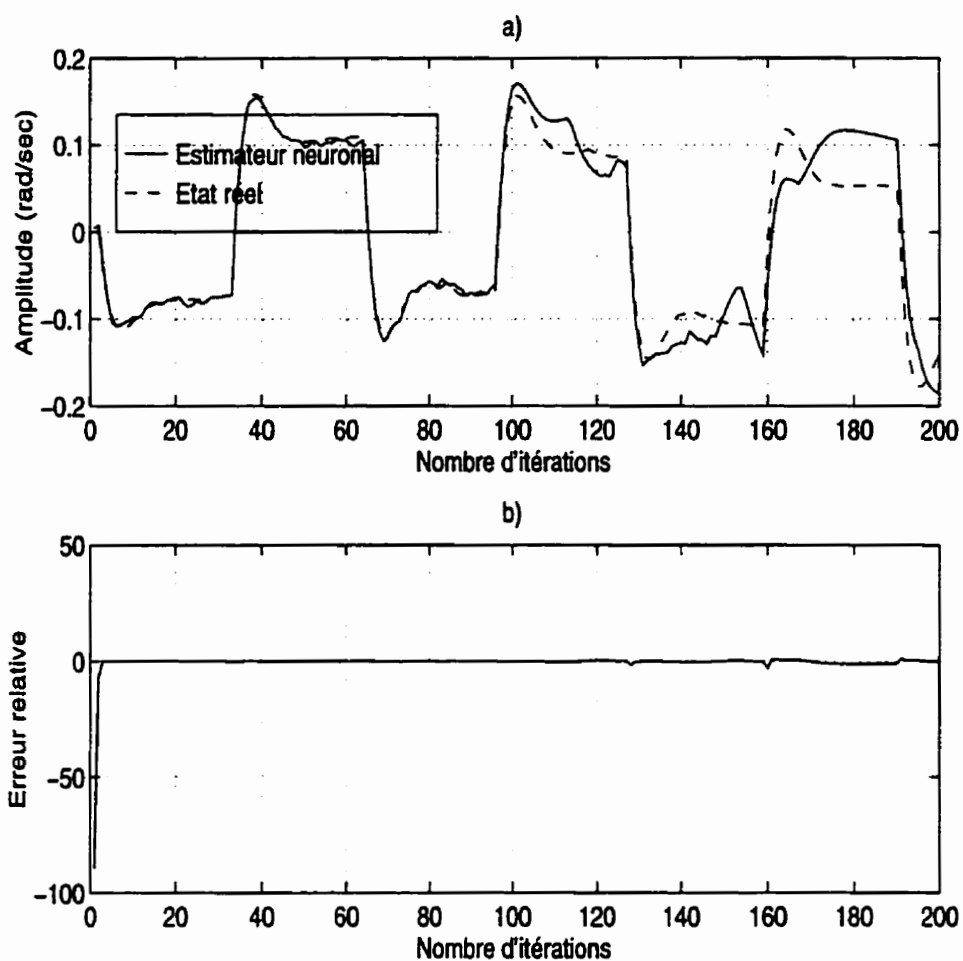


Figure 6.8: a) Etat  $x_3$  ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative.



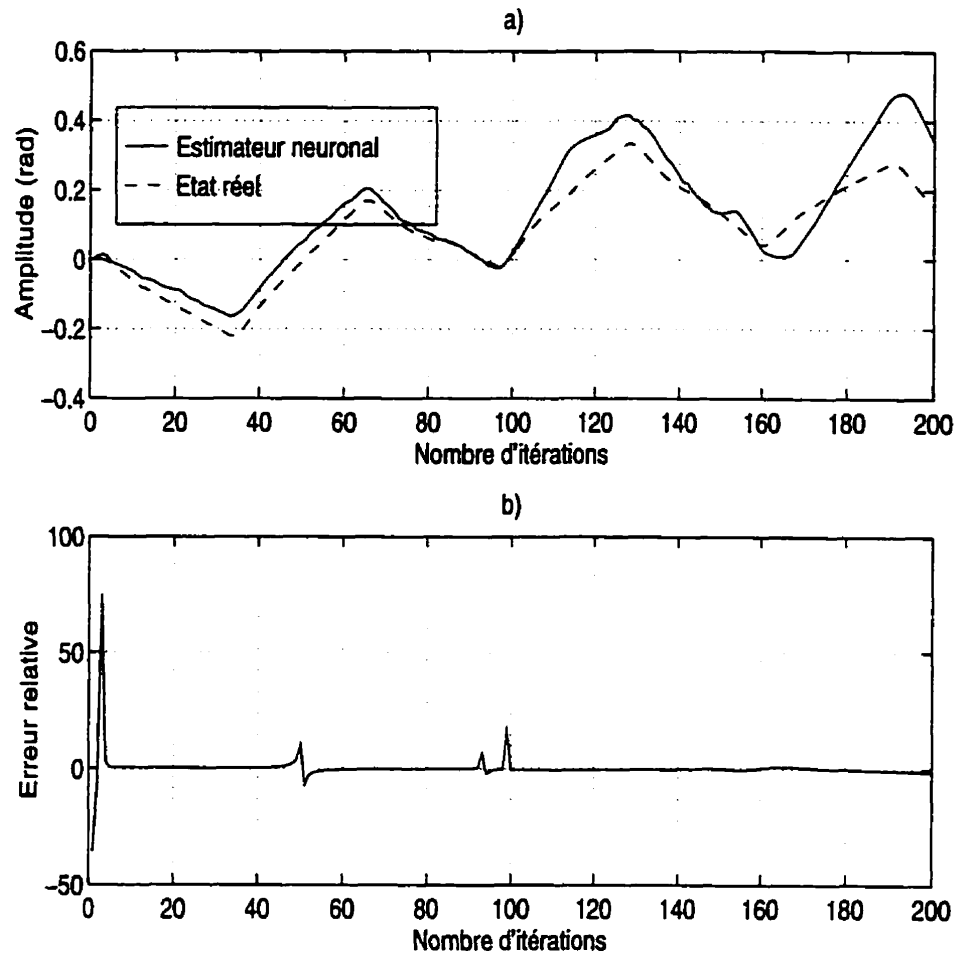
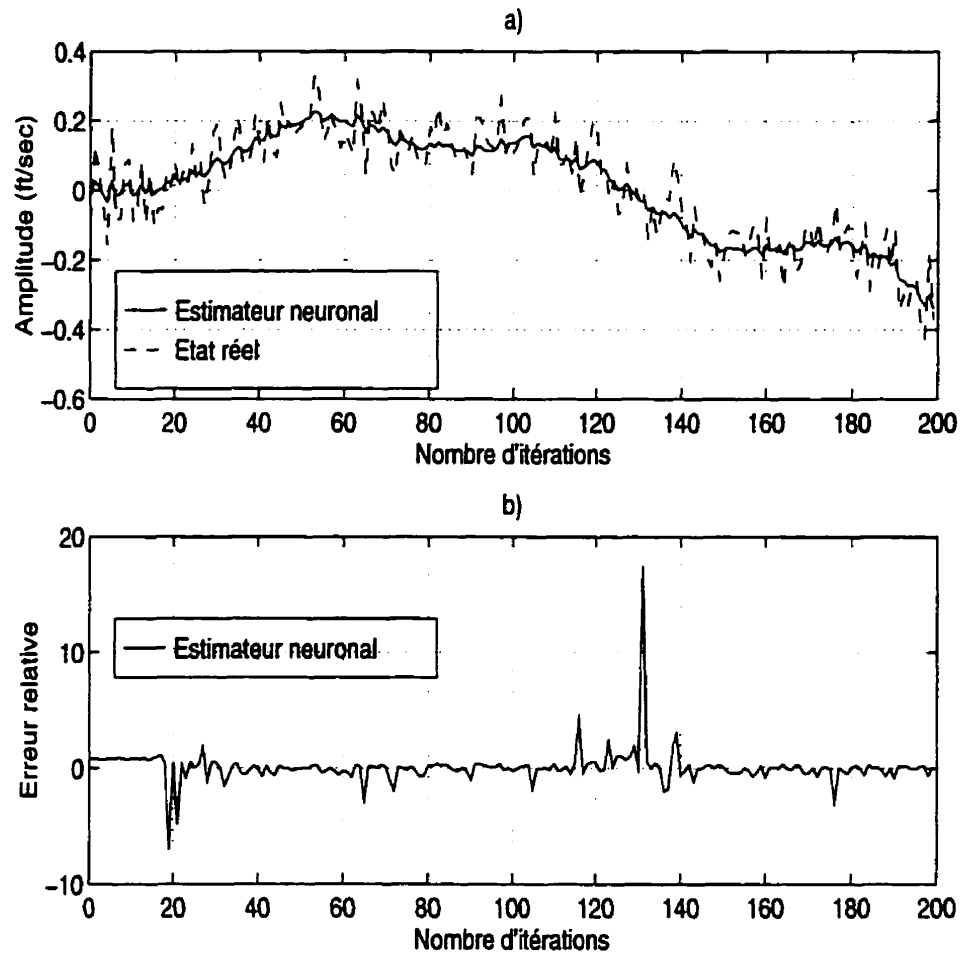


Figure 6.9: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien; b) Erreur relative.

#### 6.1.2.2 Le rapport bruit/signal à 5%

Le rapport bruit/signal est défini comme celui de la puissance du signal sur celle du bruit. Nous avons utilisé dans le cadre de ce travail le rapport des valeurs quadratiques moyennes du signal et du bruit. Si la valeur de ce rapport est de 5%, l'estimation des états bruités de l'avion par le filtre neuronal a donné les résultats des courbes des figures 6.10; 6.11; 6.12; 6.13. On peut remarquer par exemple à

la figure 6.10; que la vitesse sur l'axe de roulis estimée par le filtre neuronal ne suit plus rigoureusement sa valeur réelle bruitée. Ce constat se dégage aussi bien pour la vitesse sur l'axe de lacet (figure 6.11), la vitesse de tangage (figure 6.12) que l'angle de tangage (figure 6.13). Le réseau a tendance à estimer la valeur moyenne du signal corrompu. Ce fonctionnement est caractéristique des réseaux neuronaux artificiels. Dans certains cas, en présence du bruit, le réseau va agir comme un filtre basse-bas.



**Figure 6.10:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative.

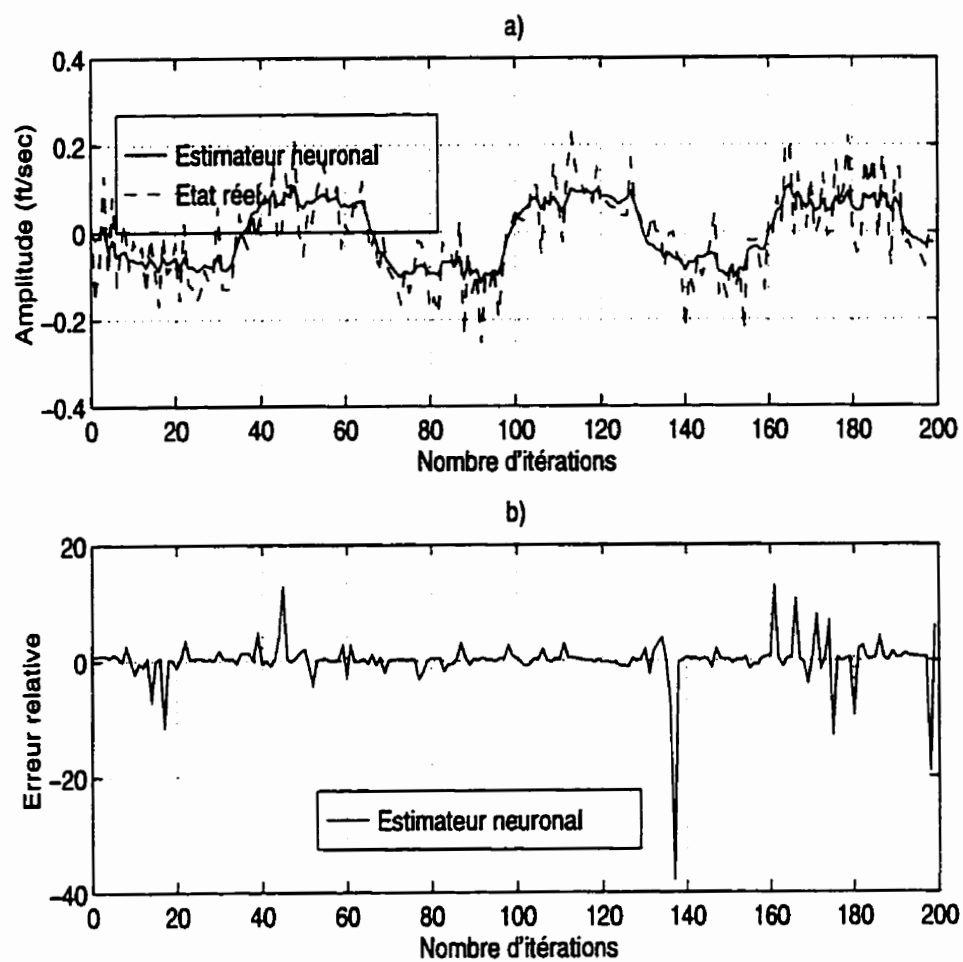
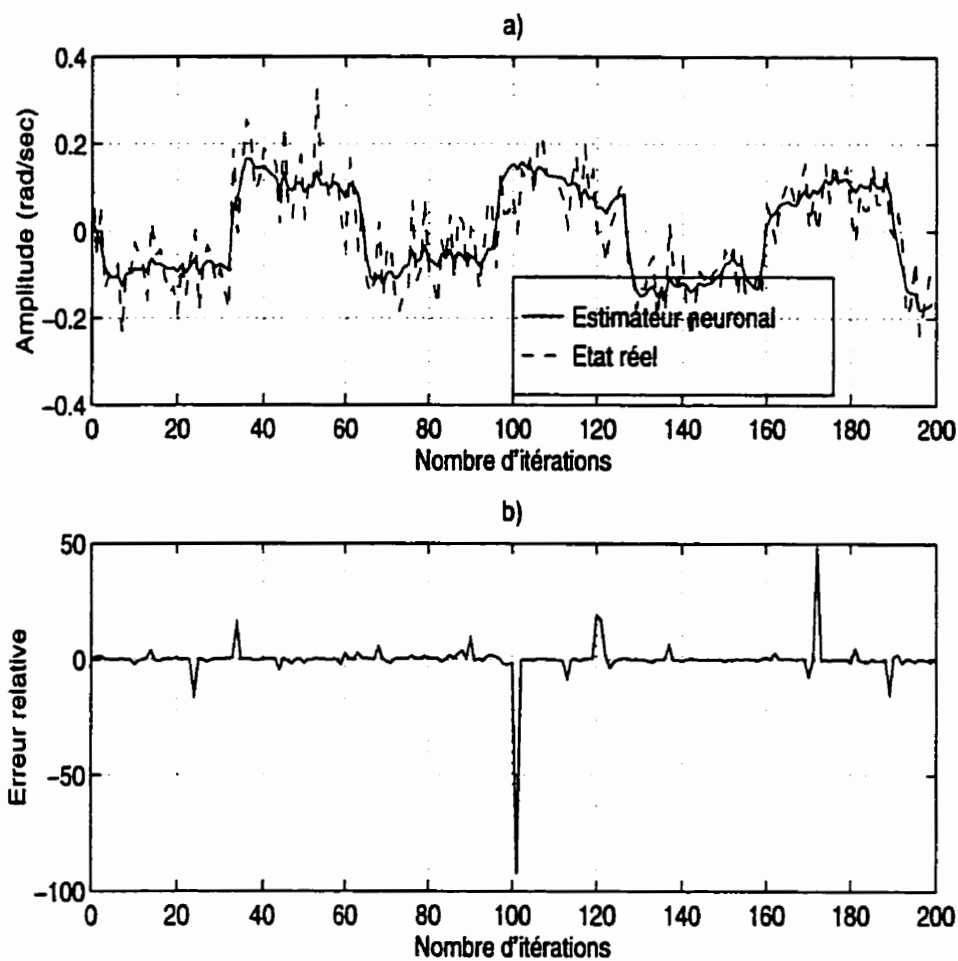
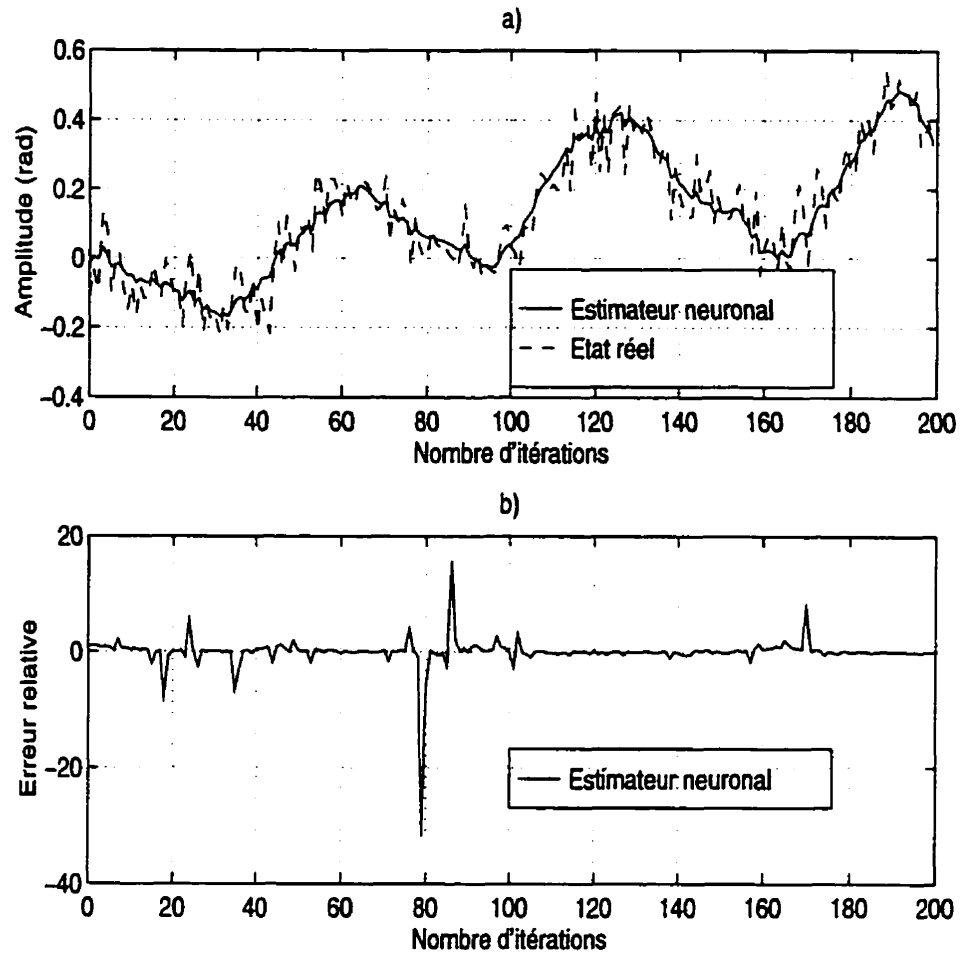


Figure 6.11: a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative.



**Figure 6.12:** a) Etat  $x_3$  ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative.



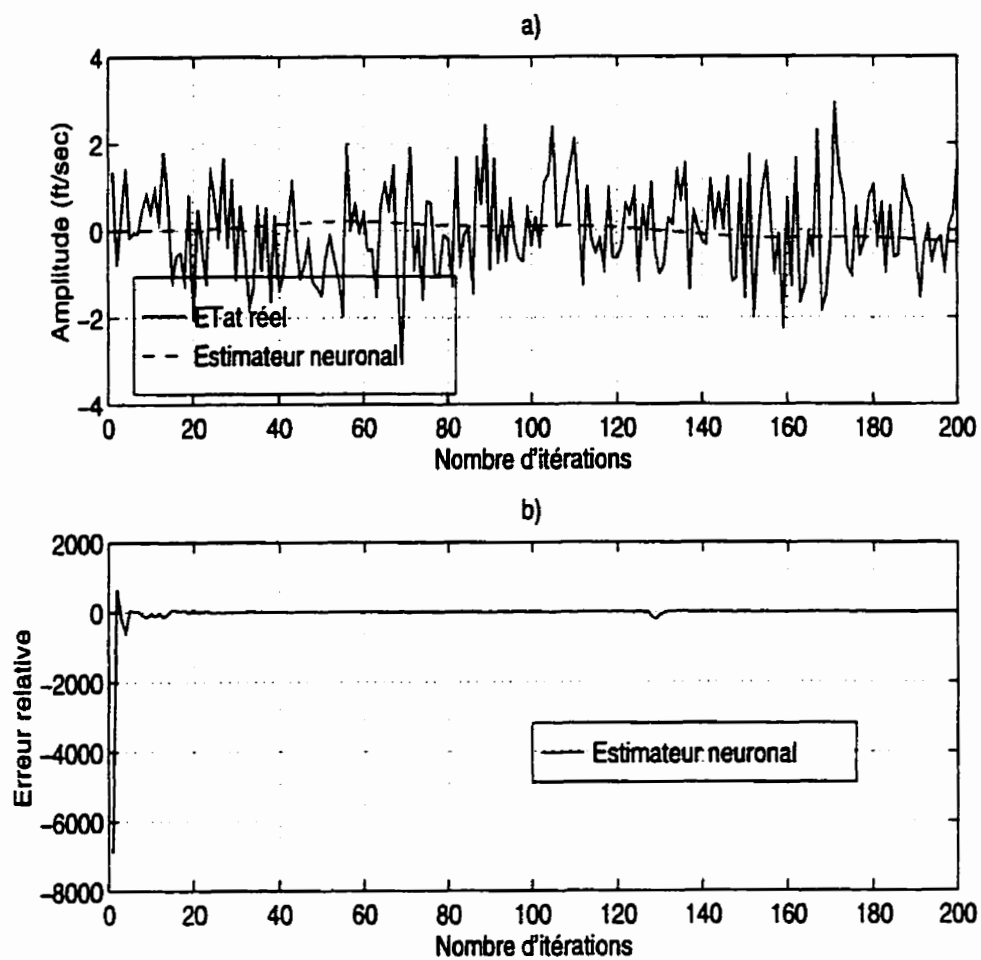
**Figure 6.13:** a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 5%; b) Erreur relative.

### 6.1.2.3 Le rapport bruit/signal à 10%

Les courbes des figures 6.14; 6.15; 6.16; 6.17 traduisent l'évolution des estimés des états de l'avion. Si le rapport signal-bruit augmente, soit à 10%, l'estimateur neuronal n'est plus capable d'estimer les états de l'avion. Les erreurs relatives deviennent plus grandes. Ce comportement du filtre peut s'expliquer par le fait que

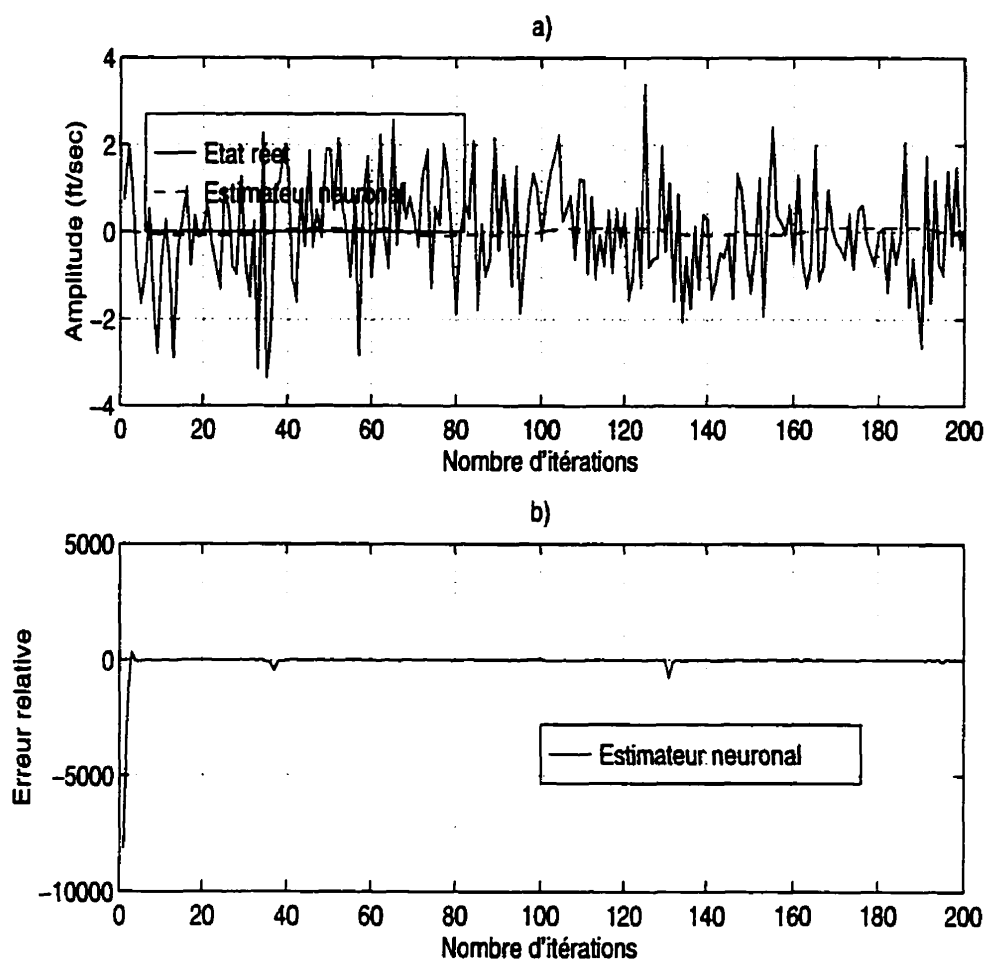
l'estimation des états par les réseaux se fait à travers la minimisation de l'erreur qui elle-même est basée sur le calcul des gradients. Si les pentes du signal sont beaucoup plus abruptes, le calcul des gradients présente des discontinuités. Le réseau évaluera ainsi de mauvais poids entraînant une estimation erronée.

En effet, le filtre neuronal ne peut pas estimer les états d'un système indépendamment du bruit qui affecte le signal. Si le bruit est dense (par rapport à l'état à estimer) le filtre produit des résultats erronés.

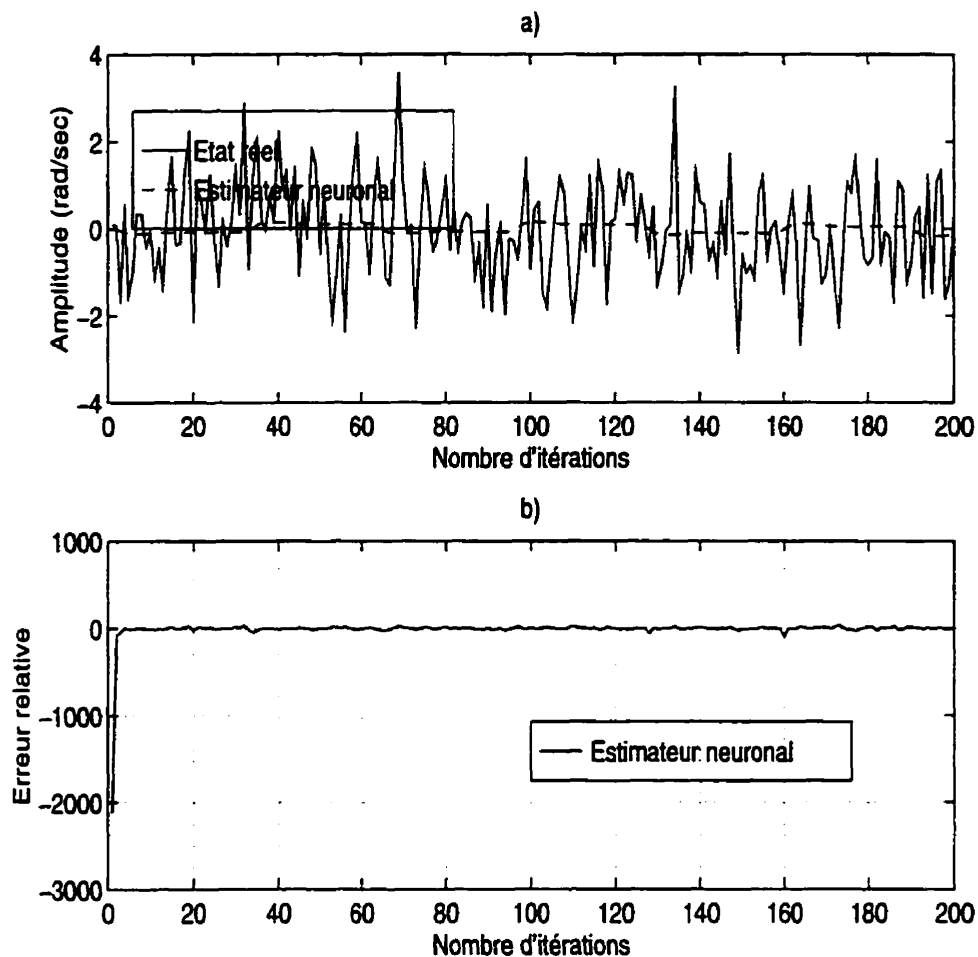


**Figure 6.14:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreur relative.





**Figure 6.15:** a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreur relative.



**Figure 6.16:** a) Etat  $x_3$  ((vitesse de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreurs relative.

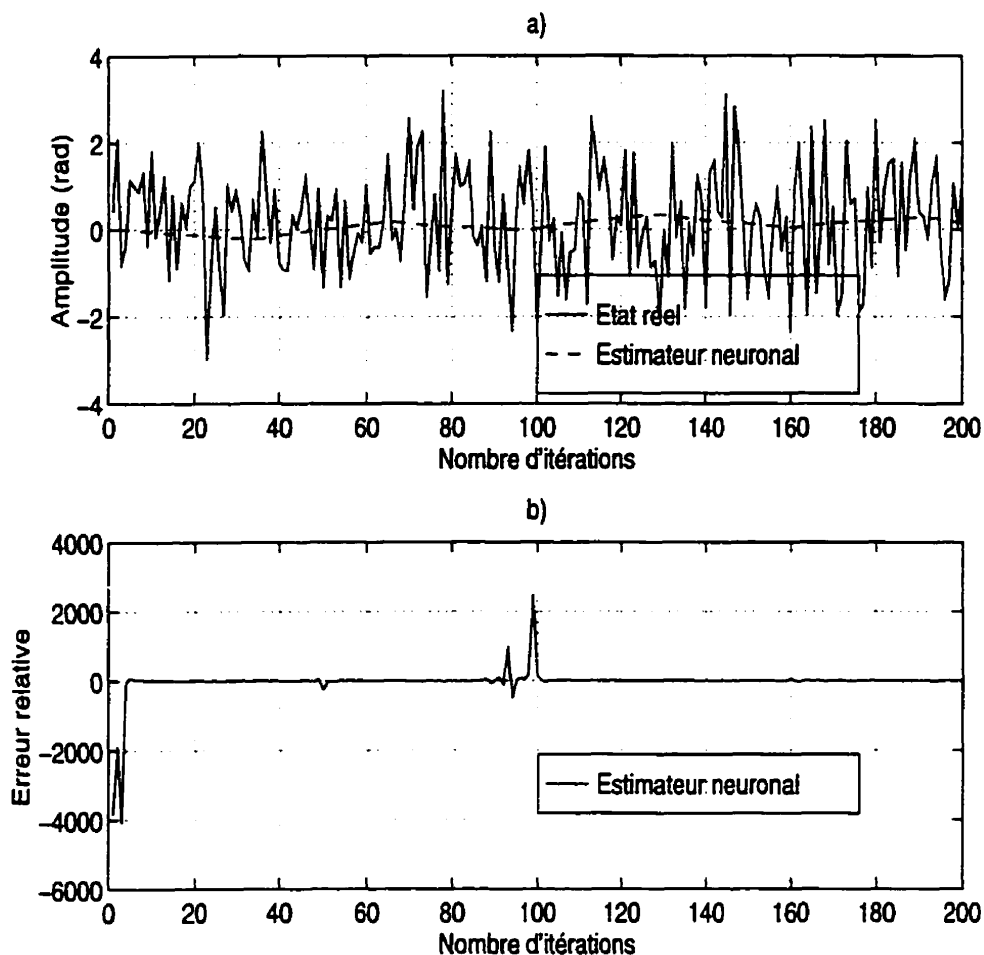


Figure 6.17: a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal dans le contexte d'un bruit non blanc et non gaussien avec rapport bruit/signal à 10%; b) Erreurs relative.

### 6.1.3 Effet de la variation des paramètres du système sur les deux estimateurs

Le filtre de Kalman dont les équations essentielles sont données par le système d'équations (2.14) est basé sur la parfaite connaissance non seulement des statis-

tiques des bruits mais aussi et surtout du modèle du système. Supposons que ce modèle soit mal connu et que le système réel soit régi par les équations (6.1) ci-dessous:

$$\begin{cases} x'_{k+1} &= \bar{F}x'_k + \bar{B}u'_k\bar{G}\omega'_k \\ y'_k &= \bar{C}x'_k + v'_k \end{cases} \quad (6.1)$$

$\omega'_k \sim (0, \bar{Q})$  et  $v'_k \sim (0, \bar{R})$  sont des bruits blancs et non corrélés et  $x'_0 \sim (E(x'_0), P_0)$ . Définissons les erreurs *a priori* et *a posteriori* entre le modèle et le système réel comme ci-dessous:

$$\tilde{x}_{k/k-1} = x'_k - x_{k/k-1} \quad (6.2)$$

$$\tilde{x}_{k/k} = x'_k - x_{k/k} \quad (6.3)$$

Cherchons à exprimer les covariances des erreurs ci-dessus:

$$\begin{aligned} \tilde{x}_{k+1/k} &= x'_{k+1} - x_{k+1/k} \\ &= \bar{F}x'_k + \bar{B}u'_k + \bar{G}\omega'_k - Fx_{k/k} - Bu_k \end{aligned} \quad (6.4)$$

Posons  $\Delta F = \bar{F} - F$ , et substituons dans (6.4)

$$\tilde{x}_{k+1/k} = F\tilde{x}_{k/k} + \Delta Fx'_k + \bar{B}u'_k - Bu_k + \bar{G}\omega'_k \quad (6.5)$$

En combinant l'équation d'état du système (6.1) et l'équation (6.5), nous obtenons:

$$\begin{bmatrix} x'_{k+1} \\ \tilde{x}_{k+1/k} \end{bmatrix} = \begin{bmatrix} \bar{F} & 0 \\ \Delta F & F \end{bmatrix} \begin{bmatrix} x'_k \\ \tilde{x}_{k/k} \end{bmatrix} + \begin{bmatrix} \bar{B} & 0 \\ \bar{B} & -B \end{bmatrix} \begin{bmatrix} u'_k \\ u_k \end{bmatrix} + \begin{bmatrix} \bar{G} \\ \bar{G} \end{bmatrix} \omega'_k \quad (6.6)$$

En définissant la matrice de corrélation comme ci-dessous:

$$\begin{bmatrix} U_k & \bar{V}_k^t \\ \bar{V}_k & \bar{S}_k \end{bmatrix} = E\left( \begin{bmatrix} x'_k \\ \tilde{x}_{k/k-1} \end{bmatrix} \begin{bmatrix} x'_k \\ \tilde{x}_{k/k-1} \end{bmatrix}^t \right) \quad (6.7)$$

et

$$\begin{bmatrix} U_k & V_k^t \\ V_k & S_k \end{bmatrix} = E \left( \begin{bmatrix} x'_k \\ \tilde{x}_{k/k} \end{bmatrix} \begin{bmatrix} x'_k \\ \tilde{x}_{k/k} \end{bmatrix}^t \right) \quad (6.8)$$

En s'inspirant de la relation (6.6) et en posant  $u_k = 0$  ;  $u'_k = 0$ , nous écrivons:

$$\begin{bmatrix} U_{k+1} & \bar{V}_{k+1}^t \\ \bar{V}_{k+1} & \bar{S}_{k+1} \end{bmatrix} = \begin{bmatrix} \bar{F} & 0 \\ \Delta F & F \end{bmatrix} \begin{bmatrix} U_k & V_k^t \\ V_k & S_k \end{bmatrix} \begin{bmatrix} \bar{F}^t & \Delta F^t \\ 0 & F^t \end{bmatrix} + \begin{bmatrix} \bar{G} \\ \bar{G} \end{bmatrix} \bar{Q} \begin{bmatrix} \bar{G}^t & \bar{G}^t \end{bmatrix} \quad (6.9)$$

ce qui donne les équations récurrentes suivantes:

$$\begin{cases} U_{k+1} = \bar{F}U_k\bar{F}^t + \bar{G}\bar{Q}\bar{G}^t \\ \bar{V}_{k+1} = \Delta F U_k \bar{F}^t + F U_k \bar{F}^t + \bar{G}\bar{Q}\bar{G}^t \\ \bar{S}_{k+1} = F S_k F^t + \bar{G}\bar{Q}\bar{G}^t + \Delta F U_k \Delta F^t \\ \quad + (F V_k \Delta F^t + \Delta F V_k^t F^t) \end{cases} \quad (6.10)$$

Il est évident que les termes ajoutés à l'équation (6.10) ( $\Delta F U_k \bar{F}^t$ ;  $\Delta F^t$  et  $F V_k \Delta F^t + \Delta F V_k^t F^t$ ) proviennent des incertitudes sur le modèle. On peut remarquer que la matrice de corrélation  $U_{k+1}$  et la matrice de corrélation croisée  $\bar{V}_{k+1}$  n'interviennent que pour la mise à jour de  $\bar{S}_{k+1}$ .  $\bar{S}_k$  étant la matrice de corrélation de l'erreur  $\tilde{x}_{k/k-1}$ , dépend par le fait même de  $U_k$ . Si  $\bar{S}_k$  est faible,  $x'_k$  et  $x_{k/k-1}$  sont proches l'un de l'autre.

Si on pose:

$$\tilde{x}_{k/k} = x'_k - x_{k/k} = x'_k - x_{k/k-1} - K_k(y'_k - C x_{k/k-1}) \quad (6.11)$$

L'utilisation de  $y'_k$  au lieu de  $y_k$  s'explique par le fait que le système réel est celui de l'équation (6.1).

$$\tilde{x}_{k/k} = \tilde{x}_{k/k-1} - K_k \bar{C} x'_k - K_k v'_k + K_k C x_{k/k-1} \quad (6.12)$$

si on pose:  $\Delta C = \bar{C} - C$

$$\tilde{x}_{k/k} = (I - K_k C) \tilde{x}_{k/k-1} - K_k \Delta C x'_k - K_k v'_k \quad (6.13)$$

Nous pouvons écrire:

$$\begin{bmatrix} x'_k \\ \tilde{x}_{k/k} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -K_k \Delta C & I - K_k C \end{bmatrix} \begin{bmatrix} x'_{k-1} \\ \tilde{x}_{k/k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ -K_k \end{bmatrix} v'_k \quad (6.14)$$

En calculant les corrélations on obtient:

$$\begin{bmatrix} U_k & \bar{V}_k^t \\ V_k & S_k \end{bmatrix} = \begin{bmatrix} I & 0 \\ -K_k \Delta C & I - K_k C \end{bmatrix} \begin{bmatrix} U_k & \bar{V}_k^t \\ \bar{V}_k & \bar{S}_k \end{bmatrix} \begin{bmatrix} I & -\Delta C^t K_k^t \\ 0 & (I - K_k C)^t \end{bmatrix} + \begin{bmatrix} 0 \\ -K_k \end{bmatrix} \bar{R} \begin{bmatrix} 0 & K_k^t \end{bmatrix} \quad (6.15)$$

Les équations récursives obtenues sont:

$$\begin{cases} U_k = U_k \\ V_k = (I - K_k C) \bar{V}_k - K_k \Delta C U_k \\ S_k = (I - K_k C) \bar{S}_k (I - K_k C)^t + K_k \bar{R} K_k^t \\ \quad + K_k \Delta C U_k \Delta C^t K_k^t \\ \quad - [(I - K_k C) \bar{V}_k \Delta C^t K_k^t \\ \quad + K_k \Delta C \bar{V}_k^t (I - K_k C)^t] \end{cases} \quad (6.16)$$

En combinant (6.4) et (6.13), nous écrivons:

$$\begin{aligned} \tilde{x}_{k+1/k} &= F(I - K_k C) \tilde{x}_{k/k-1} + (\Delta F - F K_k \Delta C) x'_k \\ &\quad + \bar{B} \bar{u}_k - B u_k + \bar{G} \omega'_k - F K_k v'_k \end{aligned} \quad (6.17)$$

La relation (6.17) montre que si  $\Delta F$  et  $\Delta C$  sont non nuls, l'erreur d'estimation est dictée par l'état  $x'_k$  du système réel. Même si l'état  $x'_k$  est borné, l'erreur  $\tilde{x}_{k/k-1}$  va croître. Il est évident de voir que même si la matrice  $F(I - K_k C)$  est stable pour  $k \rightarrow \infty$ , les erreurs de modélisation peuvent conduire le filtre à la non-convergence et à une erreur d'estimation non bornée. Par ailleurs,

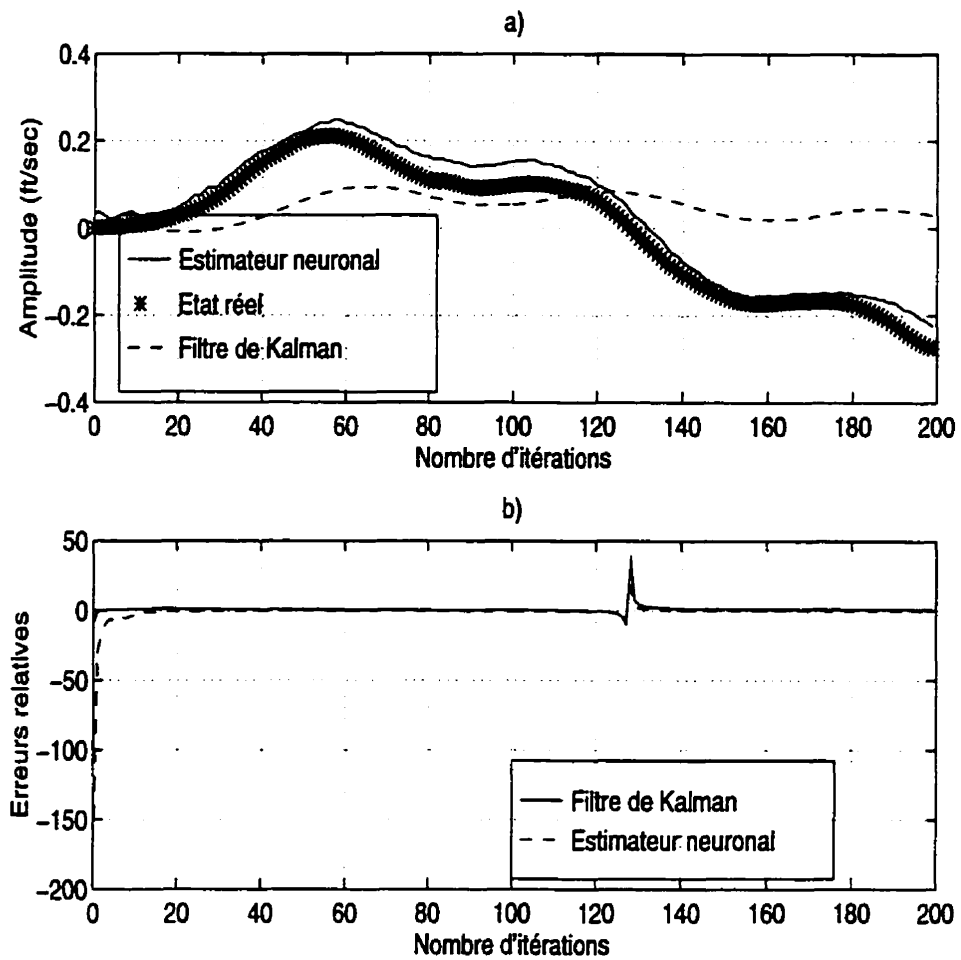
$$\begin{aligned} E(\tilde{x}_{k+1/k}) &= F(I - K_k C) E(\tilde{x}_{k/k-1}) + (\Delta F - F K_k \Delta C) E(x'_k) + \bar{B} \bar{u}_k \\ &\quad - B u_k \end{aligned} \quad (6.18)$$

même si  $E(\tilde{x}_{0/0}) = 0$  aucune garantie n'est faite pour que  $E(\tilde{x}_{k/k-1}) = 0$  en présence des incertitudes de modélisation. En somme,  $U_k$ ,  $V_k$  et  $S_k$  ne devraient pas être considérées comme des matrices de covariances mais comme des matrices de corrélation.

### 6.1.3.1 Perturbation des pôles de $F$ à 5% dans le cas d'un bruit blanc gaussien

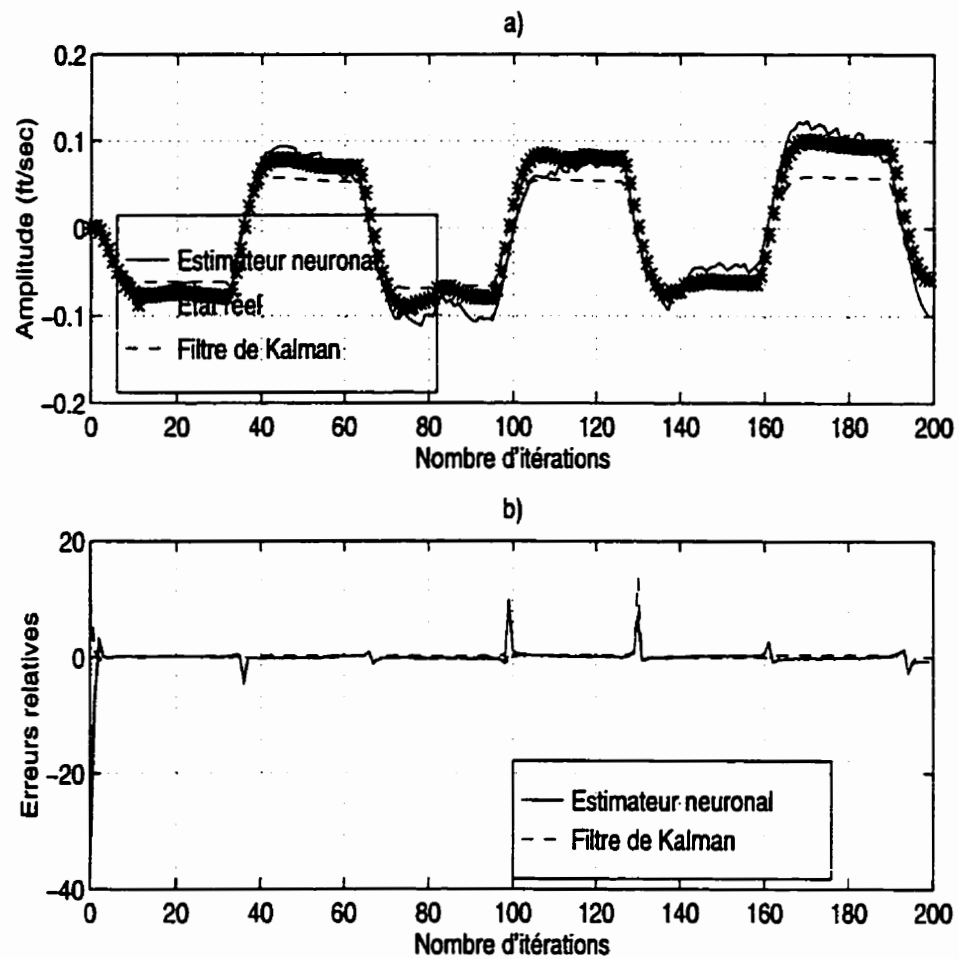
Afin de faire ressortir l'effet d'une perturbation sur la matrice d'état  $F$ , nous avons multiplié les modules des pôles de  $F$  par 5% et posé  $B = 2 * B$ . Le constat que l'on fait de cette perturbation est que les estimés du filtre de Kalman sont affectés (figures 6.18, 6.19, 6.20, 6.21). L'analyse précédente explique ce comportement. Le filtre de Kalman de base manque donc de robustesse, ce qui le rend sensible aux variations des paramètres du système. Il faut se rappeler que le filtre de Kalman minimise la variance de l'erreur et non l'erreur elle-même.

L'estimateur neuronal arrive néanmoins à maintenir les valeurs des estimés, ce qui justifie sa robustesse. Cette quasi-insensibilité aux variations des paramètres du système vient du fait que le RNA modifie ses poids en fonction de l'erreur  $y_k - \hat{y}_k$  dans le but de la minimiser. Si  $F$  est perturbée, le RNA cherchera toujours à faire tendre  $\hat{y}_k$  vers  $y_k$ . Il reste cependant à voir si cette insensibilité persiste lorsque la matrice d'état  $F$  du système est suffisamment perturbée. Pour analyser ce cas, nous perturberons à 10% les pôles du système en boucle ouverte au paragraphe suivant.

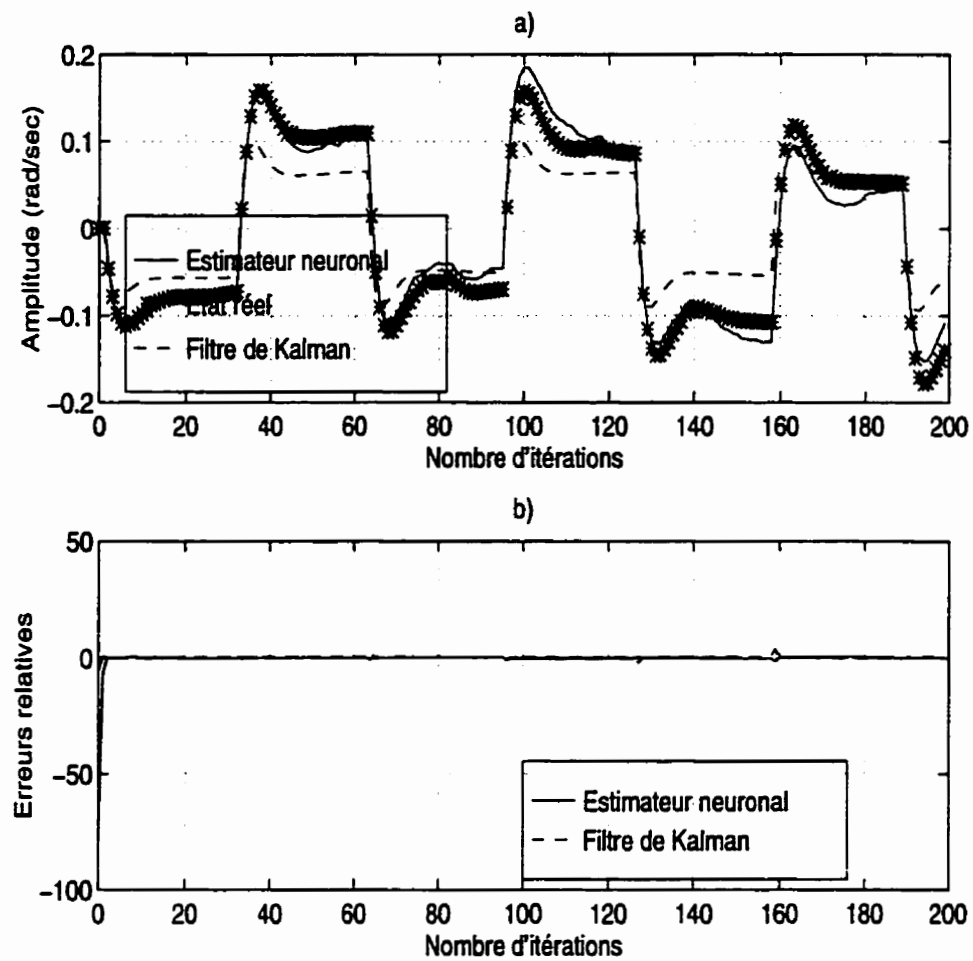


**Figure 6.18:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives.

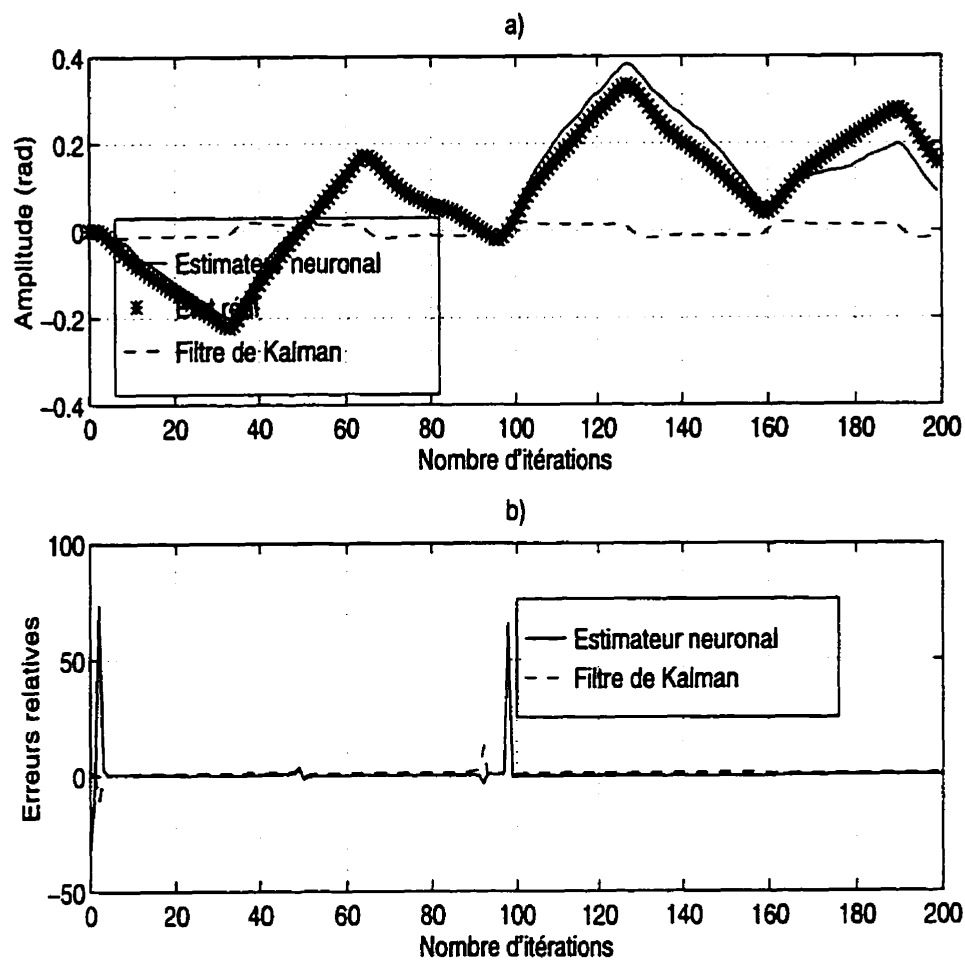




**Figure 6.19:** a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives.



**Figure 6.20:** a) Etat  $x_3$  (vitesse de tangage) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives.

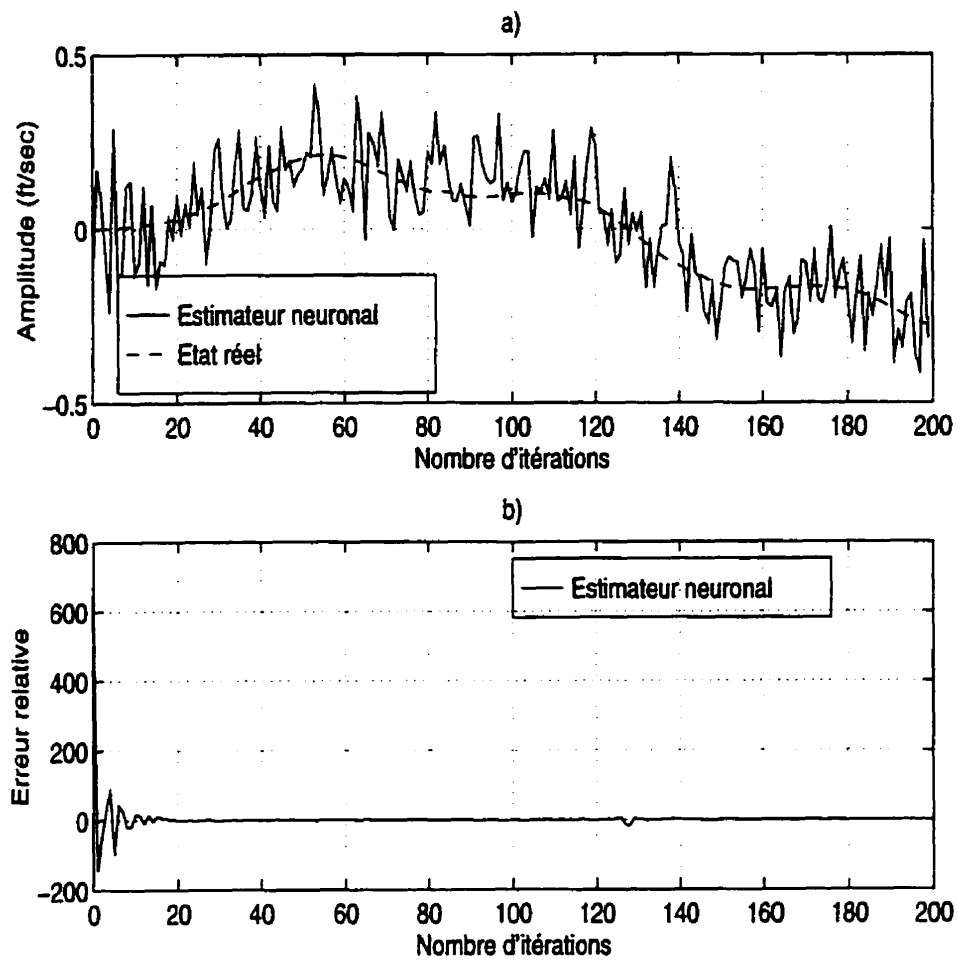


**Figure 6.21:** a) Etat  $x_4$  (angle de tangage) estimé par le filtre neuronal et le filtre de Kalman quand les pôles de  $F$  sont perturbés à 5%; b) Erreurs relatives.

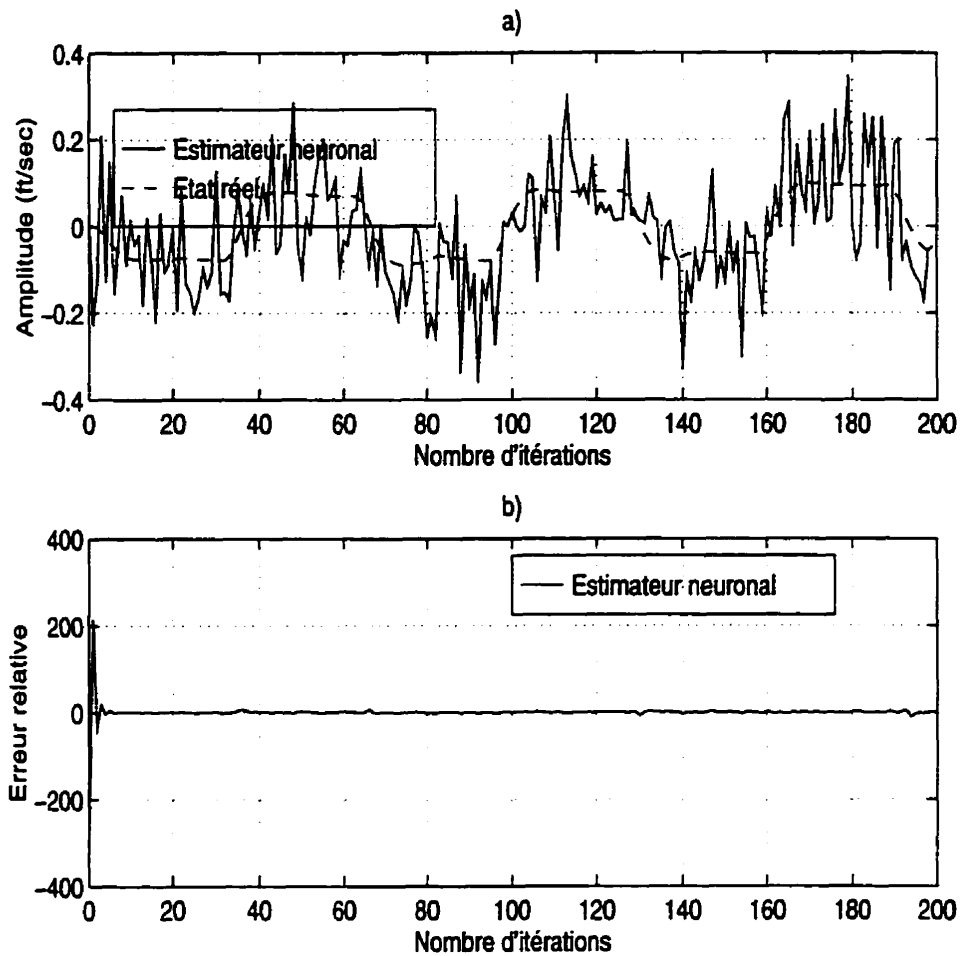
### **6.1.3.2 Perturbation à 10% des pôles du système dans le cas d'un bruit blanc gaussien**

Des figures 6.22; 6.23; 6.24; 6.25, on peut facilement remarquer que la robustesse de l'estimateur neuronal est fonction du degré de perturbation de la matrice d'état  $F$ . A une perturbation de 10% des pôles du système, l'estimateur donne de mauvais résultats. Aussi bien la vitesse sur l'axe de roulis, la vitesse sur l'axe de lacet, la vitesse de tangage que l'angle de tangage. Ceci signifie que les variables d'état sont mal estimés lorsque les pôles du système sont largement modifiés.

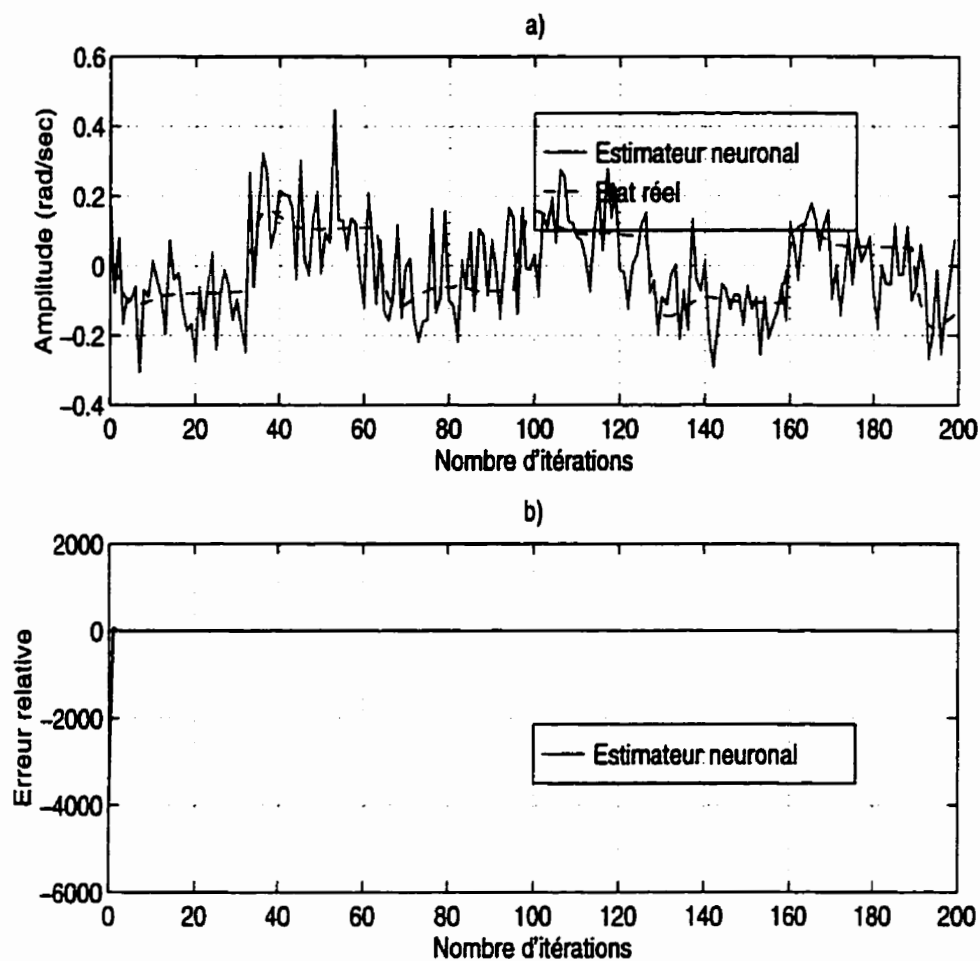
Pour ce système en étude, le filtre n'arrive à supporter que des variations de 5%; la robustesse du filtre n'est donc pas indépendante du degré de variation des pôles.



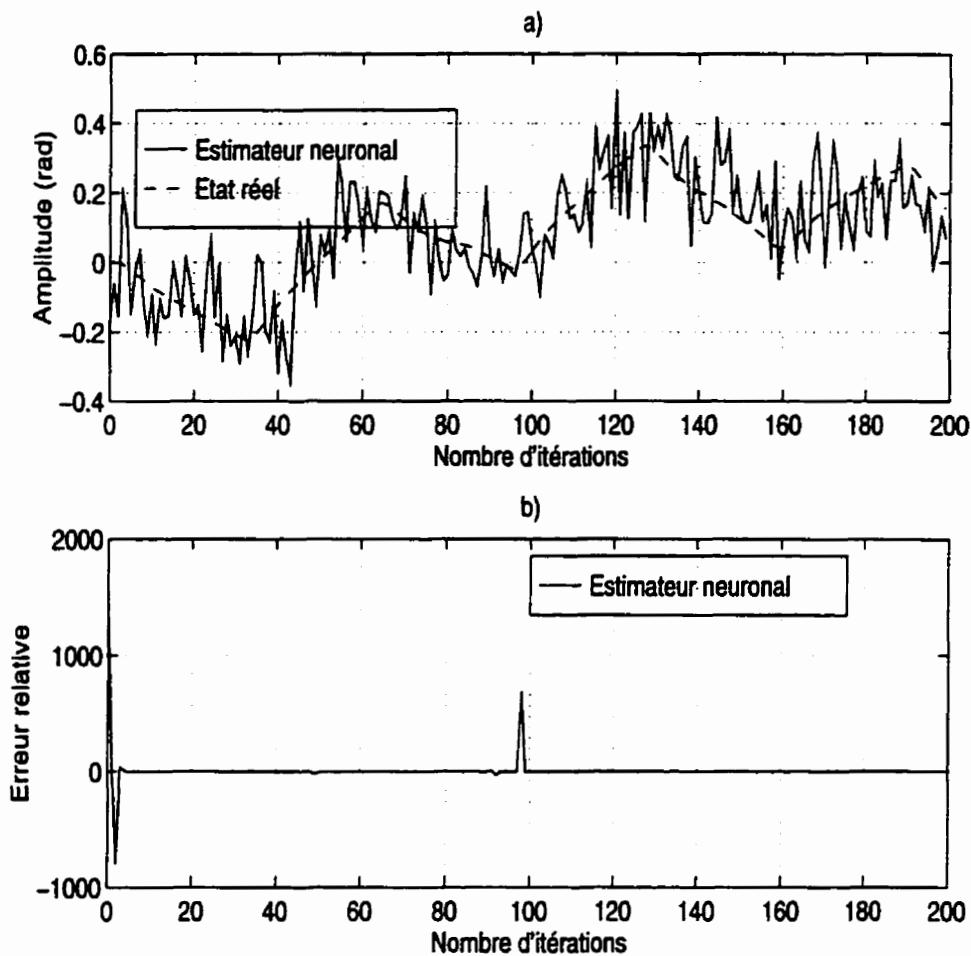
**Figure 6.22:** a) Etat  $x_1$  (vitesse sur l'axe de roulis) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative.



**Figure 6.23:** a) Etat  $x_2$  (vitesse sur l'axe de lacet) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative.



**Figure 6.24:** a) Etat  $x_3$  (vitesse de tangage) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative.



**Figure 6.25:** a) Etat  $x_1$  (angle de tangage) estimé par le filtre neuronal quand les pôles de  $F$  sont perturbés à 10%; b) Erreur relative.

## Conclusion

L'analyse comparative des résultats donnés par les deux filtres; le filtre de Kalman et l'estimateur neuronal, dans un environnement stochastique a montré qu'ils ont un comportement semblable. L'estimateur neuronal présente, entre autres, une



quasi insensibilité aux perturbations des paramètres du système. Toutefois, cette insensibilité est limitée au cas où le rapport signal-bruit est moins élevé. Aussi avons-nous constaté que l'estimateur neuronal donne un résultat acceptable pour un rapport bruit/signal à 5%. Ce résultat se détériore lorsque ce rapport est de 10%. L'estimateur neuronal a un avantage de pouvoir estimer les états des systèmes en l'absence des statistiques des bruits qui les affectent. Cependant, il est nécessaire que le degré de bruit soit faible si l'on veut obtenir de meilleurs résultats.

## Chapitre 7

# ANALYSE COMPARATIVE DU FILTRE DE KALMAN ET DU PRÉDICTEUR NEURONAL

### Introduction

Ce chapitre est, lui aussi, une bonne base de comparaison entre le filtre de Kalman et le prédicteur neuronal . Cette comparaison sera axée sur les comportements des deux prédicteurs dans les cas suivant:

- les bruits sont blancs et gaussiens et corrompent le système très faiblement
- le rapport bruit/signal est de 5% (bruits sont blancs et gaussiens)
- le rapport bruit/signal est de 10% (bruits sont blancs et gaussiens)

Par ailleurs, on s'intéressera aussi aux séquences de mesures incomplètes. En effet, on reprendra les trois cas précédents en supposant que les mesures des apports

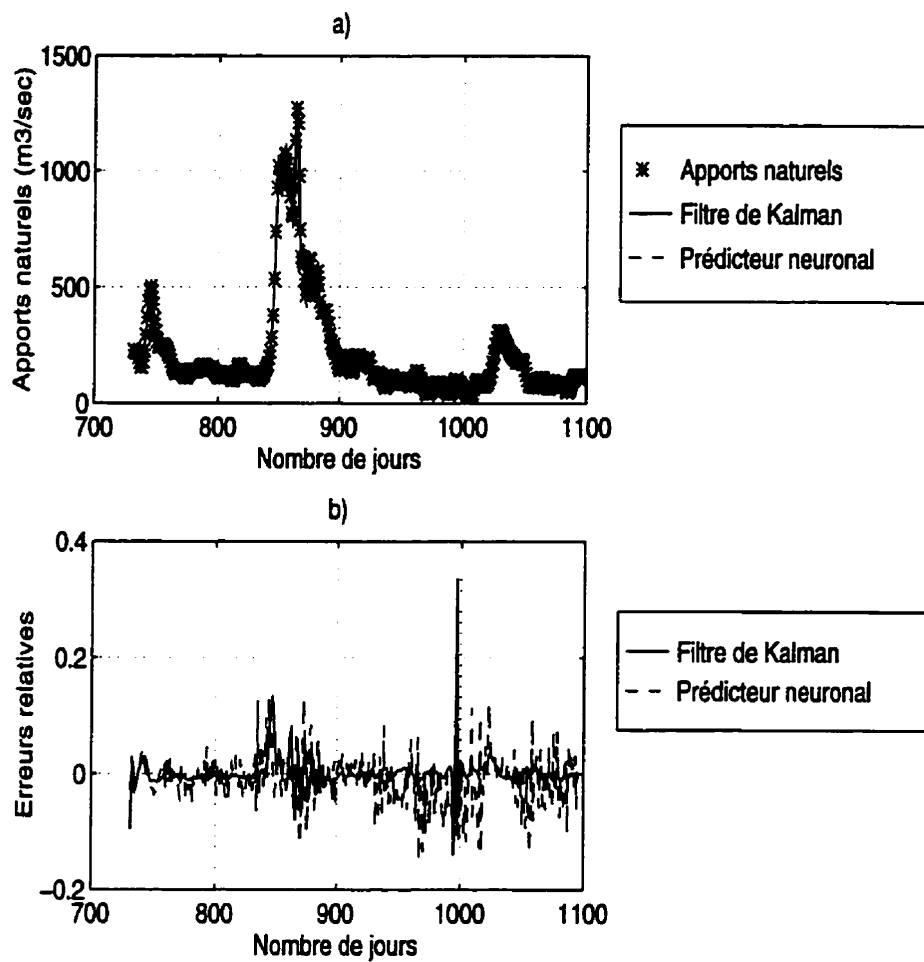
naturels sont incomplètes et qu'il faudra prédire les mesures manquantes.

## 7.1 MESURES COMPLÈTES

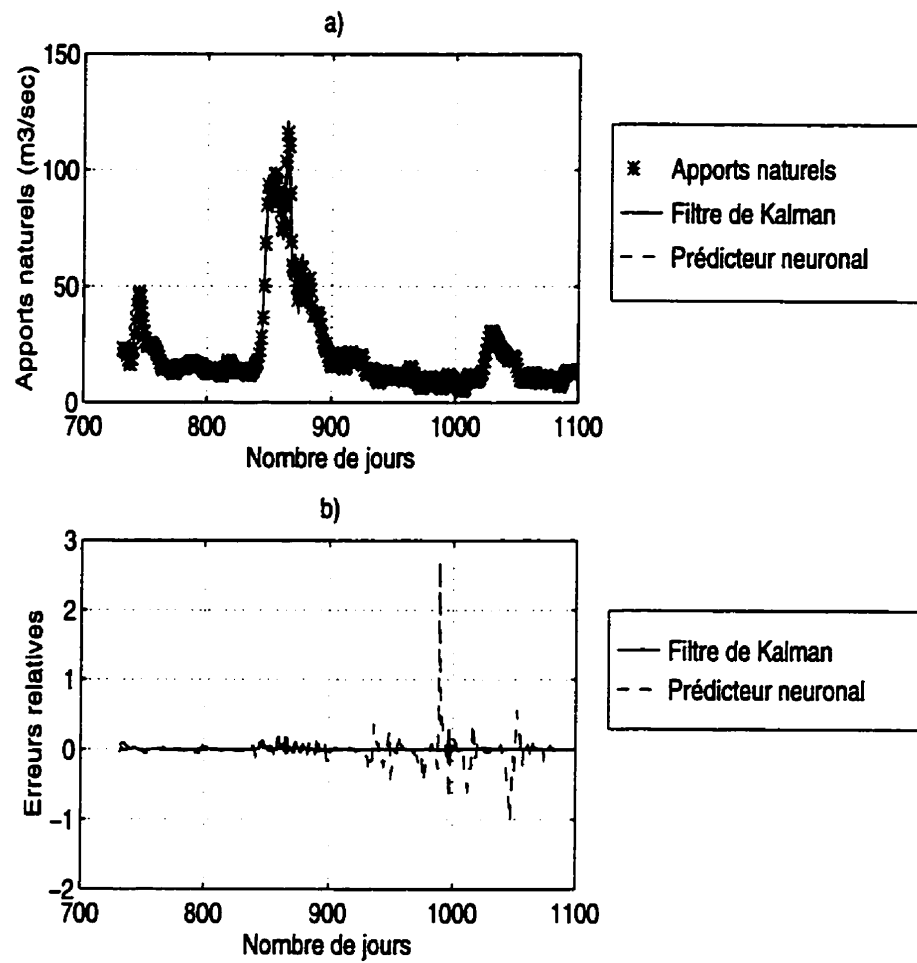
### 7.1.1 Les bruits sont blancs et gaussiens

Nous supposons que toutes les observations sont parfaitement connues et qu'en outre, les bruits qui affectent le système (trois réservoirs) sont blancs gaussiens dont les moyennes et les covariances ont été identifiées par la méthode de Assaf et Quick, (1991). Les figures 7.1;7.2;7.3 montrent les prédictions des apports naturels des trois réservoirs par filtre de Kalman et par prédicteur neuronal. Les deux prédicteurs aboutissent très sensiblement aux mêmes résultats. Les erreurs relatives sont relativement faibles. Le filtre de Kalman étant optimal dans ce cas de bruits blancs gaussiens. Le fait que ses résultats soient presque identiques à ceux fournis par le prédicteur neuronal, permet d'avancer que le filtre neuronal n'est pas loin de l'optimalité.

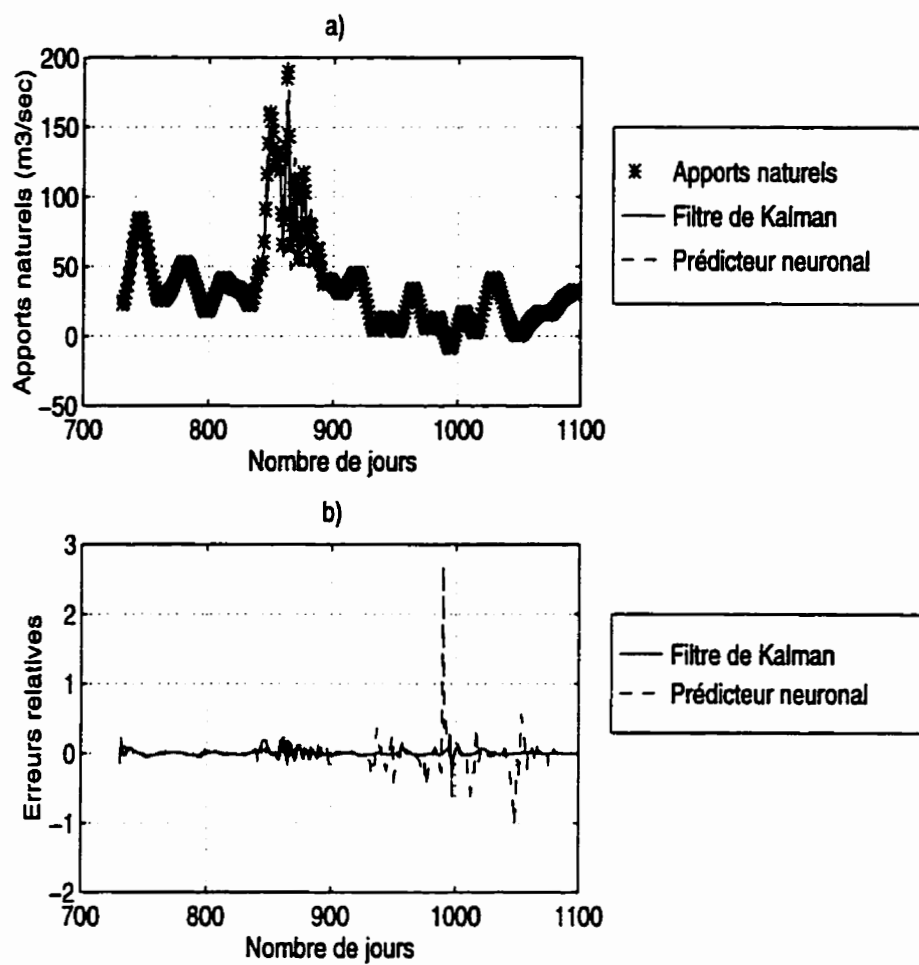
Les données des apports naturels des réservoirs de l'Hydro-Québec ont été traitées; c'est-à-dire presque dépourvues d'une grande proportion de bruit. Si ces données venaient davantage à être corrompues avec un rapport bruit/signal élevé; le filtre neuronal aura-t-il le même comportement?



**Figure 7.1:** a) Prédiction à observations complètes; b) Erreurs moyennes de prédiction (Baskatong).



**Figure 7.2:** a) Prédiction à observations complètes; b) Erreurs relatives; rapport bruit/signal à 1% (Maniwaki).



**Figure 7.3:** a) Prédiction à observations complètes; b) Erreurs relatives; rapport bruit/signal à 1% (Cabonga).

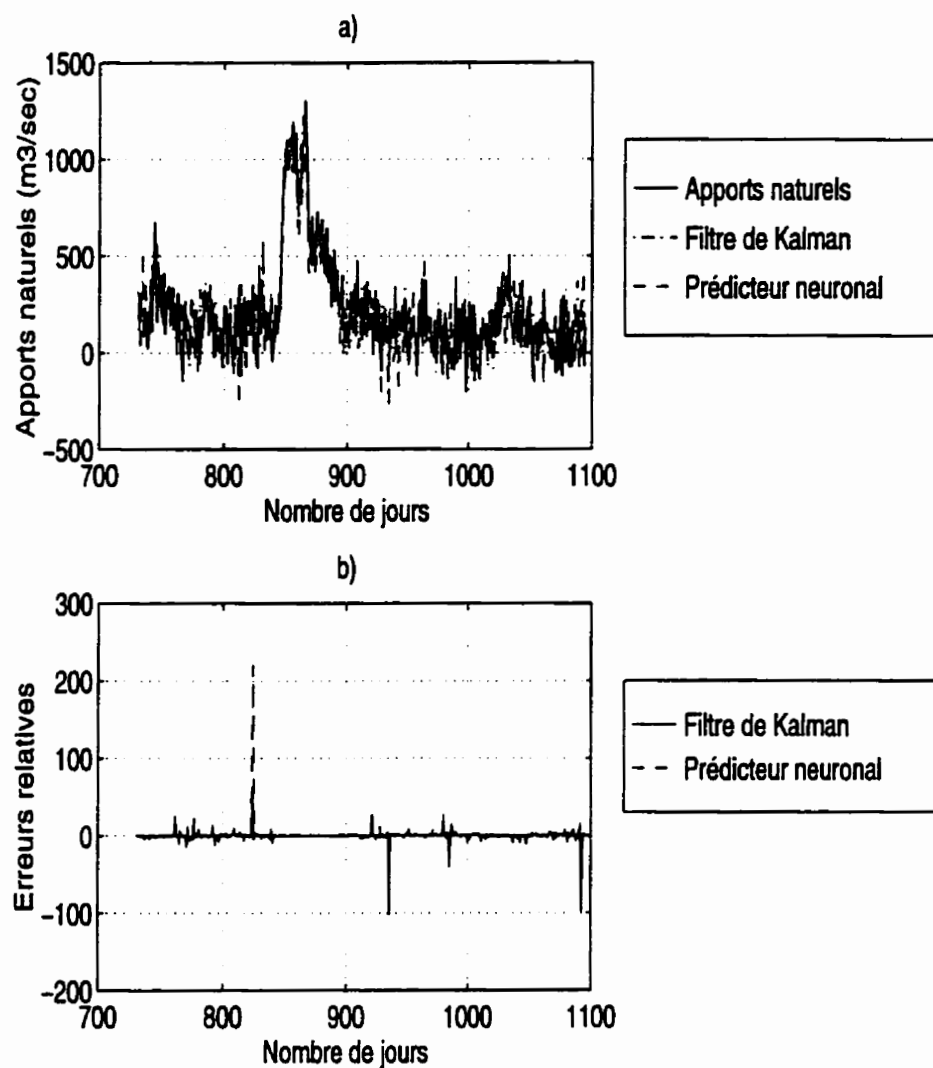
### 7.1.2 Le rapport bruit/signal à 5%

Dans ce cas de données bruitées, le modèle des réservoirs est toujours celui développé au chapitre 5 avec trois mesures passées pour chaque réservoir et neuf variables d'état. Les paramètres estimés sont ceux du tableau 7.1. Notons que chaque fois que l'intensité du bruit est changée, un nouvel apprentissage du prédicteur neuronal est fait. Ici les apports naturels ont été corrompus et le rapport bruit/signal est de 5%. Les courbes de prédiction des apports naturels des trois réservoirs des figures 7.6;7.5;7.4 montrent que la prédiction du filtre neuronal n'est plus aussi bonne que lors du précédent cas. Les erreurs relatives sont relativement grandes. Ce phénomène provient du fait qu'au lieu de prédire les valeurs bruitées, le réseau a tendance à se débarrasser du bruit. Le filtre de Kalman, quant à lui, donne néanmoins une bonne prédiction bien que ses erreurs relatives soient aussi un peu plus élevées.

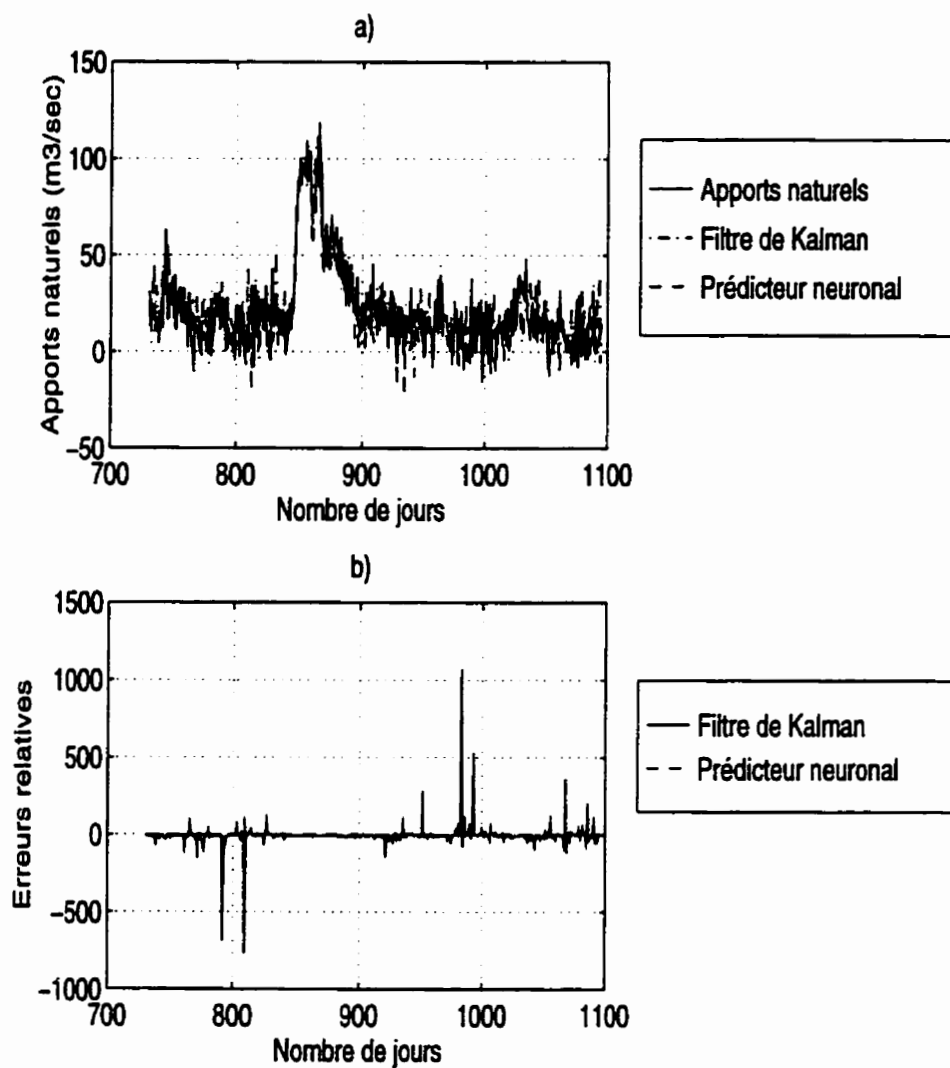
**Tableau 7.1:** Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu. Mesures bruitées avec un rapport signal-bruit de 5%

$a_1$	$b_1$	$c_1$
46.0934	5.9444	4.0687
$a_2$	$b_2$	$c_2$
-14.2545	-1.7076	-1.2952
$a_3$	$b_3$	$c_3$
-31.6802	-4.3734	-2.8491
$a_4$	$b_4$	$c_4$
-3.4706	1.5836	-0.3096
$a_5$	$b_5$	$c_5$
6.1916	-0.8956	0.5515
$a_6$	$b_6$	$c_6$
-2.4034	0.3267	-0.2141
$a_7$	$b_7$	$c_7$
-501.4488	-67.9221	-44.2508
$a_8$	$b_8$	$c_8$
151.3569	20.9042	13.7690
$a_9$	$b_9$	$c_9$
357.9959	48.3638	32.1938

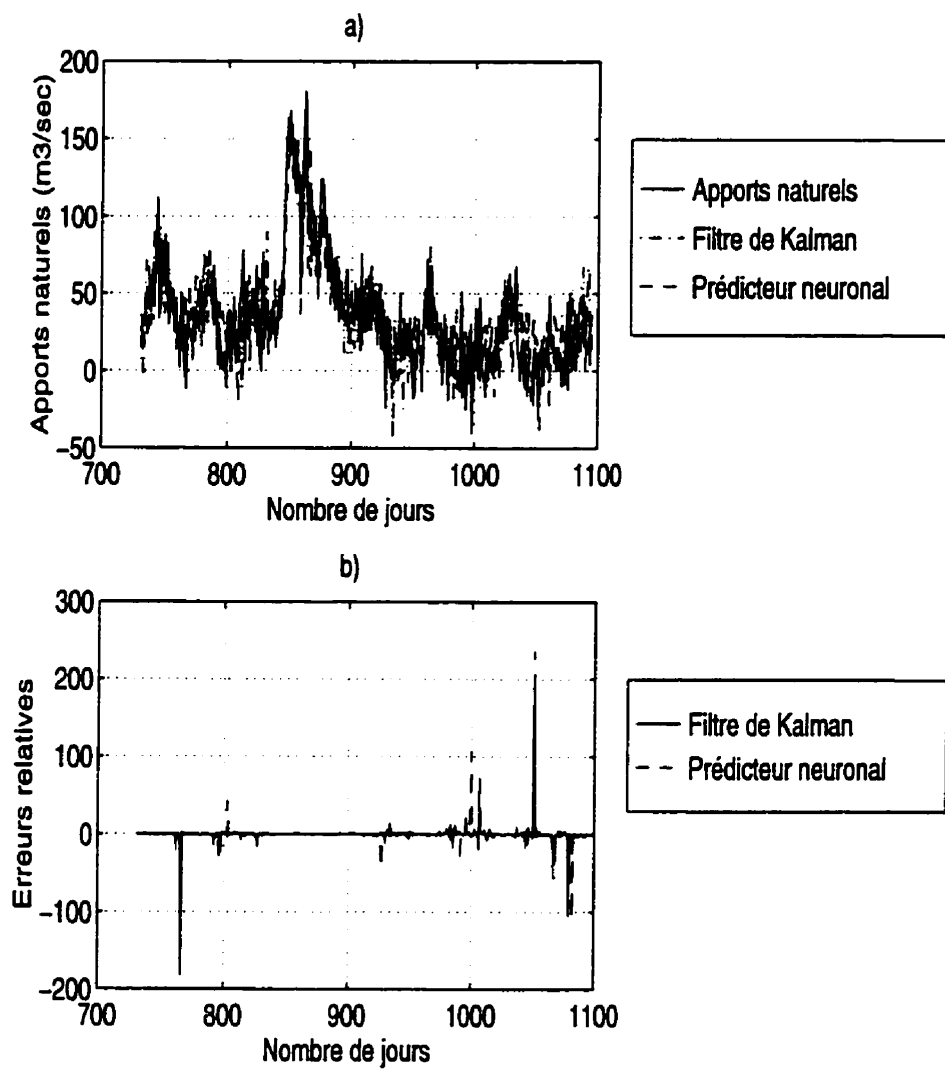




**Figure 7.4:** a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Baskatong).



**Figure 7.5:** a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Maniwaki).



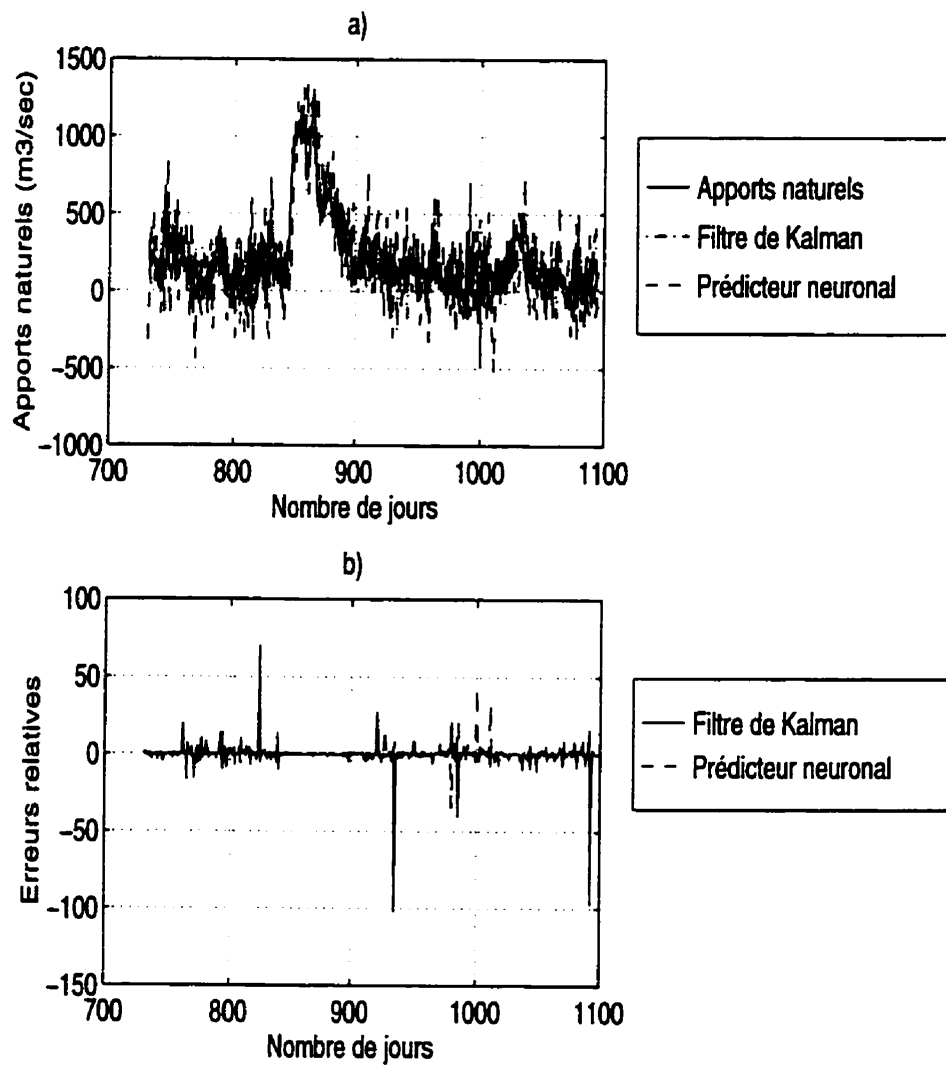
**Figure 7.6:** a) Prédiction à observations complètes; rapport bruit/signal à 5%; b) Erreurs relatives (Cabonga).

### 7.1.3 Le rapport bruit/signal à 10%

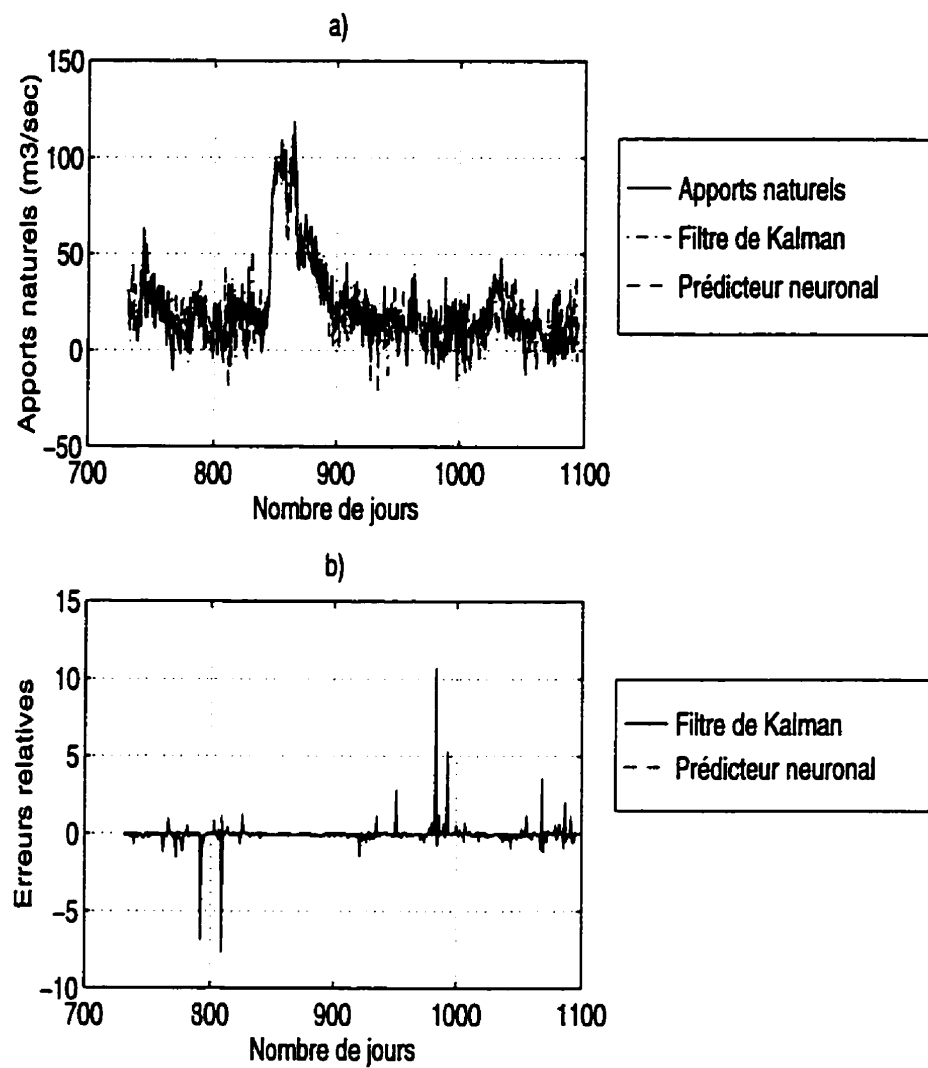
Dans ce cas de données bruitées, le modèle des réservoirs est toujours celui développé au chapitre 5 avec trois mesures passées pour chaque réservoir et neuf variables d'état, Les paramètres estimés sont ceux du tableau 7.2. L'on peut remarquer aux figures 7.9;7.8;7.7 qu'à ce degré de corruption des données des réservoirs, les deux prédicteurs ne donnent plus de bons résultats. Le prédicteur neuronal produit des erreurs moyennes élevées. On peut dire que le filtre neuronal n'est plus utilisable. Ceci donne un indice quant à la performance du prédicteur neuronal. Si le rapport bruit/signal est élevé, il n'est pas utile de prédire avec le filtre neuronal.

**Tableau 7.2:** Paramètres du modèle des réservoirs (modèle basé sur trois mesures passées) identifiés par le filtre de Kalman étendu. Mesures bruitées avec un rapport bruit/bruit de 10%

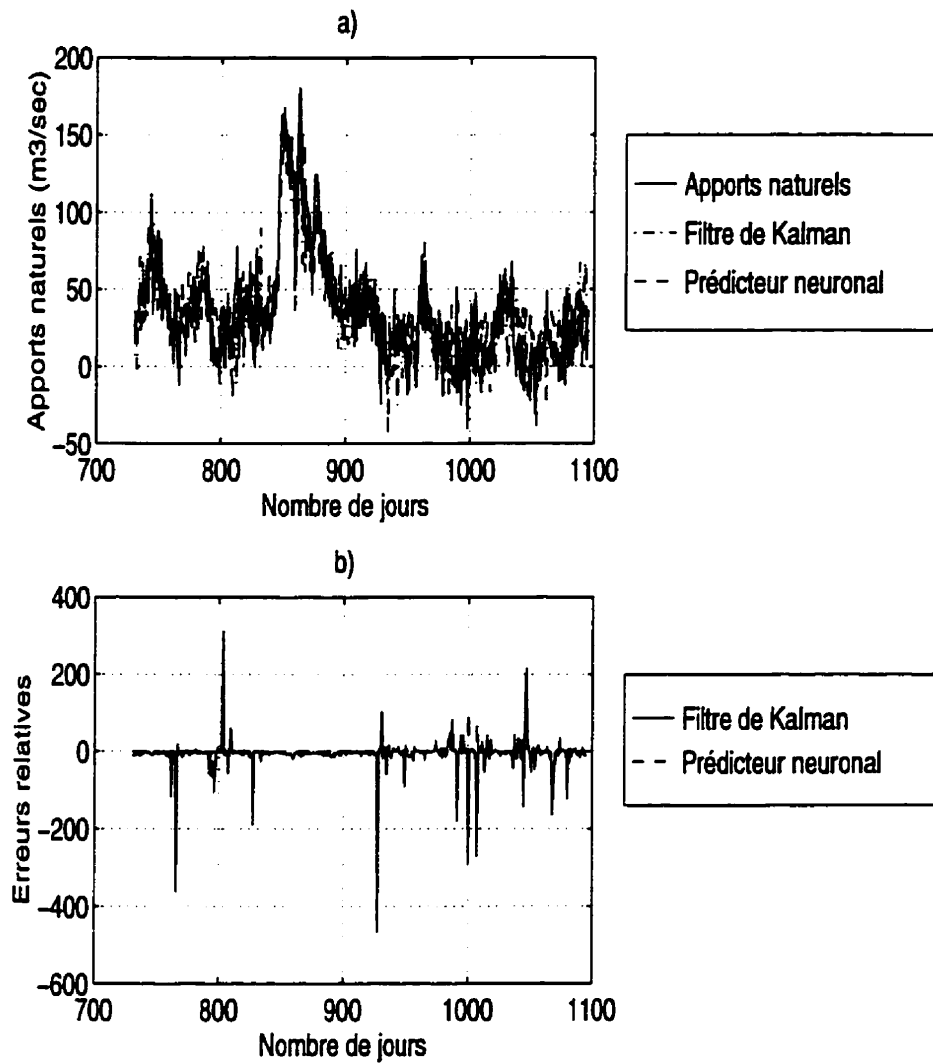
$a_1$	$b_1$	$c_1$
44.7606	5.7217	3.9455
$a_2$	$b_2$	$c_2$
-8.7947	-0.8922	-0.8022
$a_3$	$b_3$	$c_3$
-37.0818	-5.1530	-3.3338
$a_4$	$b_4$	$c_4$
-4.3568	1.4557	-0.3893
$a_5$	$b_5$	$c_5$
5.6405	-0.9810	0.5016
$a_6$	$b_6$	$c_6$
-1.3644	0.4817	-0.1204
$a_7$	$b_7$	$c_7$
-486.1854	-65.37.20	-42.8408
$a_8$	$b_8$	$c_8$
90.8219	11.8539	8.3031
$a_9$	$b_9$	$c_9$
416.6840	56.8314	37.4594



**Figure 7.7:** a) Prédiction à observations complètes; rapport bruit/signal à 10%; b) Erreurs relatives (Baskatong).



**Figure 7.8:** a) Prédiction à observations complètes; rapport signal-bruit à 10%; b) Erreurs relatives.(Maniwaki)



**Figure 7.9:** a) Prédiction à observations complètes; rapport signal-bruit à 10%; b) Erreurs relatives (Cabonga).



## 7.2 MESURES INCOMPLÈTES

Pour créer cette situation d'informations incomplètes, nous avons sur un an (entre le 731<sup>ème</sup> et le 1095<sup>ème</sup> jour) présenté aux deux prédicteurs des données du 731<sup>ème</sup> au 831<sup>ème</sup> jour. Les données entre le 831<sup>ème</sup> et le 1031<sup>ème</sup> jour sont supposées inconnues. Les données entre le 1032<sup>ème</sup> et le 1095<sup>ème</sup> jour sont elles aussi présentées aux deux prédicteurs.

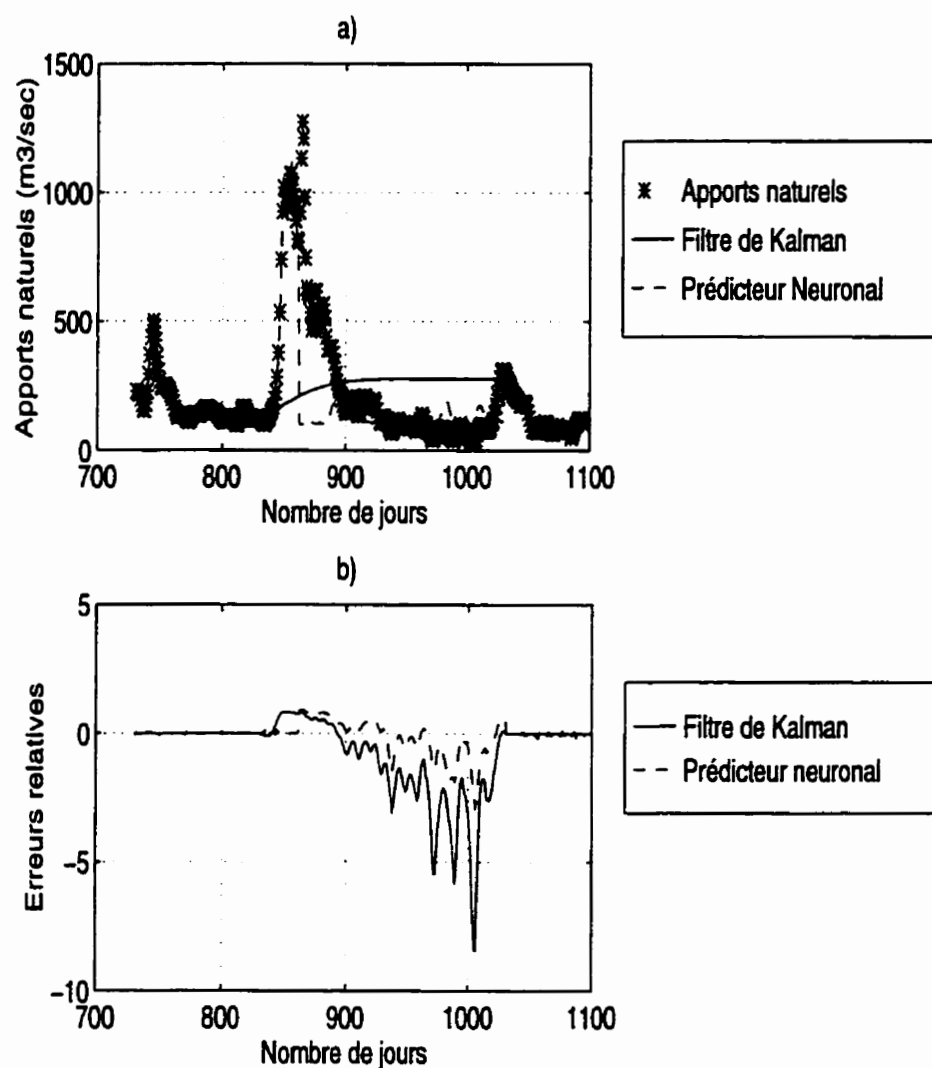
Les résultats observés sont ceux apparaissant aux figures 7.10;7.11;7.12 pour les trois réservoirs. Nous remarquons que les deux prédicteurs donnent de meilleurs résultats lorsque les mesures sont présentées. C'est presque le même résultat qu'on a obtenu au paragraphe précédent lorsque les observations sont complètes. Dans le cas où les mesures sont manquantes, le filtre de Kalman n'est plus capable de prédire de bonnes valeurs des apports naturels des réservoirs tel que montré au tableau 7.3. Ce comportement n'est pas surprenant dans la mesure où le filtre de Kalman minimise la variance de l'erreur. Si l'innovation est absente, le filtre ne peut plus s'adapter. Par ailleurs, le filtre de Kalman est un prédicteur à un pas dans le temps. Le prédicteur neuronal réussit néanmoins à produire véritablement une dizaine de bonnes valeurs de prédiction des apports naturels avant de voir les valeurs suivantes se détériorer au fur et à mesure que la prédiction se poursuit. Cette prédiction vient nécessairement du fait que le filtre neuronal a d'abord procédé à un apprentissage dont les données s'apparentent à celles qu'il faut prédire. En effet à partir de la figure 4.7; le manque de mesures se traduit par l'ouverture de  $y_k$ . Les entrées du réseaux à cet instant  $k$  sont  $y_{1k-1}; y_{1k-2}; y_{1k-3}; y_{2k-1}; y_{2k-2}; y_{2k-3}; y_{3k-1}; y_{3k-2}; y_{3k-3}; -\hat{y}_{1k}; -\hat{y}_{2k} -\hat{y}_{3k}$ . La présence de ces entrées permettra au réseau de donner quelques bonnes valeurs de prédiction en l'absence des mesures. Il faut

rappeler ici que le réseau de neurones a une capacité d'apprentissage de l'espace dans le temps qui lui permet de fournir, en l'absence de toute mesure, des valeurs proches à celles supposées estimées.

Comme cela a été dit précédemment, les apports naturels de l'Hydro-Québec sont traitées c'est-à-dire faiblement corrompues. Si le rapport bruit/signal était élevé, le réseau prédirait-il toujours de bonnes valeurs en l'absence des mesures?

**Tableau 7.3:** Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures (les mesures sont traitées)

Jours	Cabonga			Maniwaki			Baskatong		
	Kalman	Reseau	Reel	Kalman	Reseau	Reel	Kalman	Reseau	Reel
831	24.79	24.69	24.30	13.27	13.50	13.20	117.80	120.10	116.90
832	24.83	23.05	22.90	12.77	12.21	12.50	111.86	105.70	108.50
833	25.81	22.25	22.60	12.53	11.91	12.40	109.25	102.20	107.20
834	27.17	23.50	23.30	12.69	12.73	11.80	111.11	111.50	100.30
835	28.43	24.70	25.10	13.28	11.05	12.50	117.69	92.60	108.70
836	29.39	27.51	28.00	14.12	14.12	13.30	127.13	127.10	117.50
837	30.07	30.75	32.20	14.97	13.97	13.70	136.73	125.40	122.20
838	30.56	35.78	37.50	15.64	13.91	14.40	144.26	124.80	130.50
839	30.98	41.70	44.20	16.02	15.45	15.70	148.53	142.10	144.30
840	31.35	49.82	50.40	16.13	17.21	16.70	149.82	161.90	156.60
841	31.69	54.20	49.70	16.10	17.84	17.80	149.46	168.90	168.20
842	32.01	45.40	47.90	16.07	18.72	20.50	149.11	178.90	198.30
843	32.34	47.18	52.60	16.15	23.51	22.80	150.03	232.70	224.70
844	32.74	60.16	67.70	16.39	24.28	27.90	152.75	241.40	282.90
845	33.22	84.97	90.90	16.77	34.54	36.40	157.01	356.80	377.70
846	33.77	111.94	116.30	17.22	46.24	50.30	162.05	488.50	534.50
847	34.35	143.44	138.20	17.66	68.68	68.60	167.05	740.50	739.80
848	34.89	153.98	158.20	18.05	88.17	85.10	171.41	960.10	925.90
849	35.36	172.68	159.80	18.36	98.24	93.50	174.89	1073.30	1021.00
850	35.74	143.70	155.50	18.60	94.15	92.90	177.59	1027.40	1013.10



**Figure 7.10:** a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Baskatong).

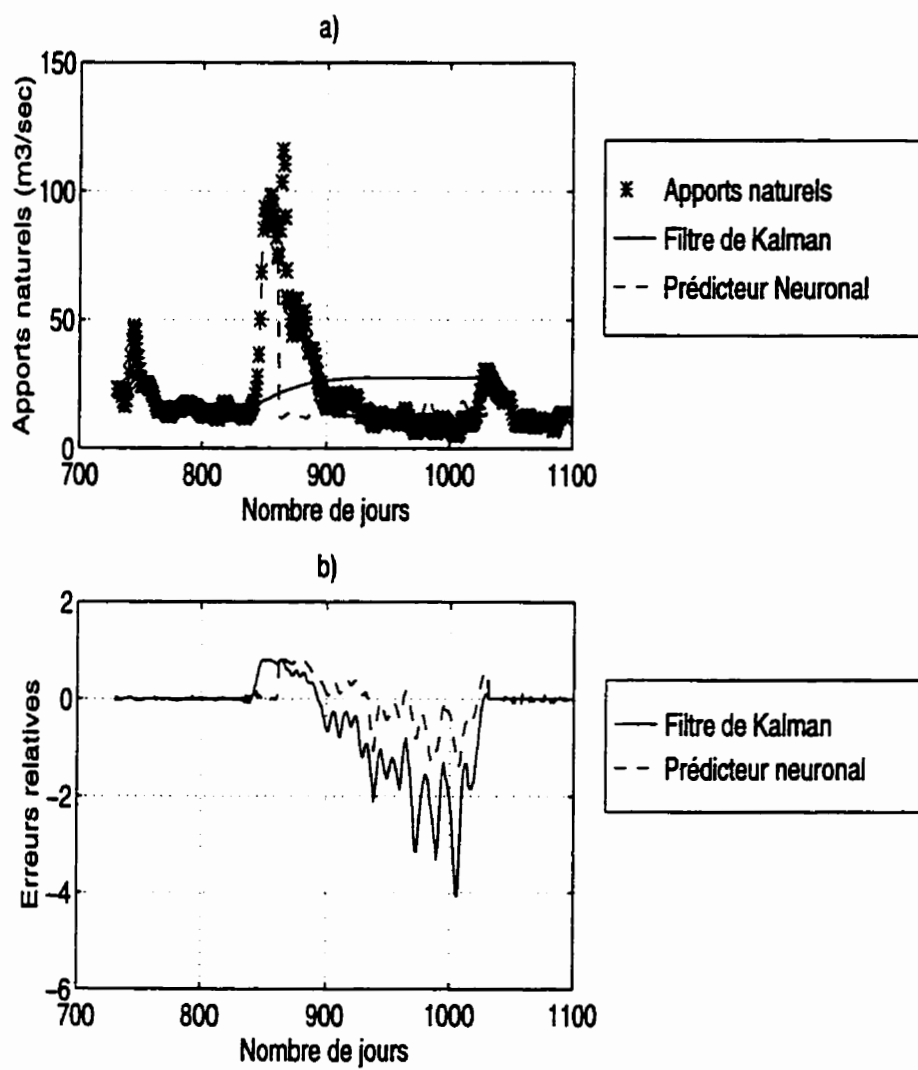
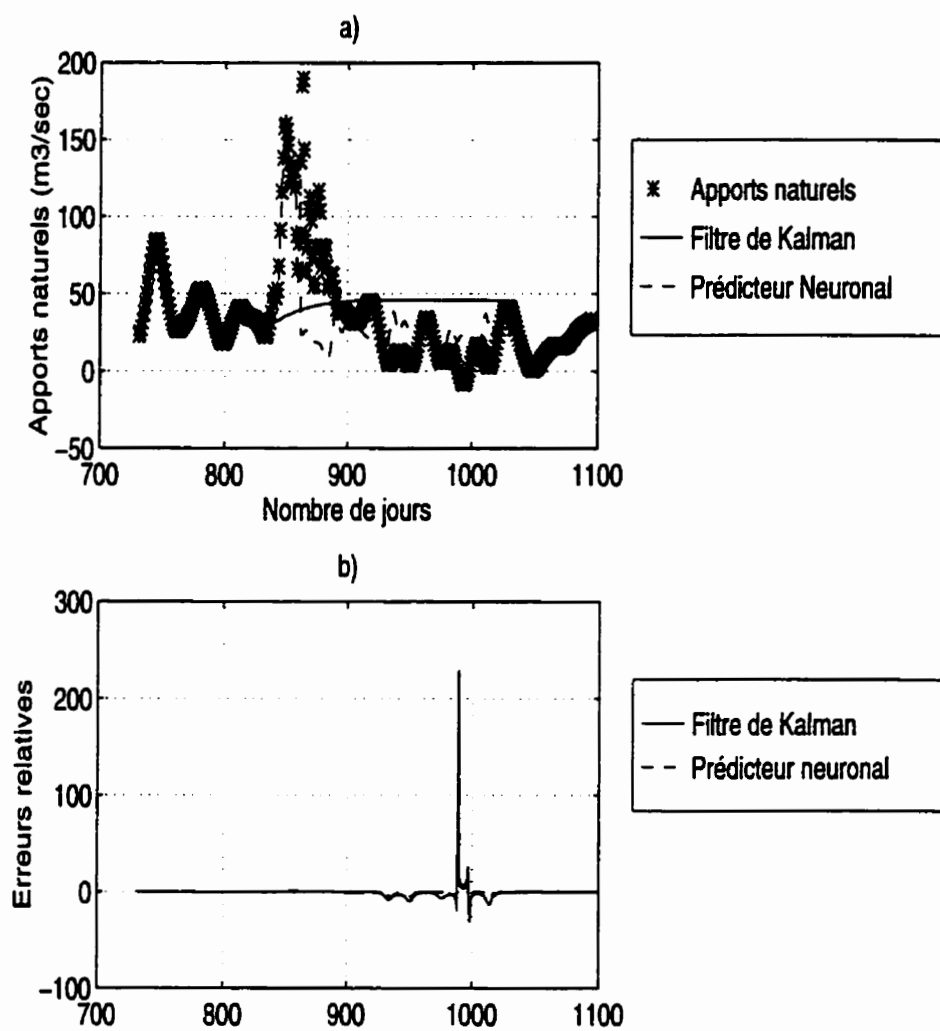


Figure 7.11: a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Maniwaki).



**Figure 7.12:** a) Prédiction à observations incomplètes; b) Erreurs relatives; rapport bruit/signal à 1% (Cabonga).

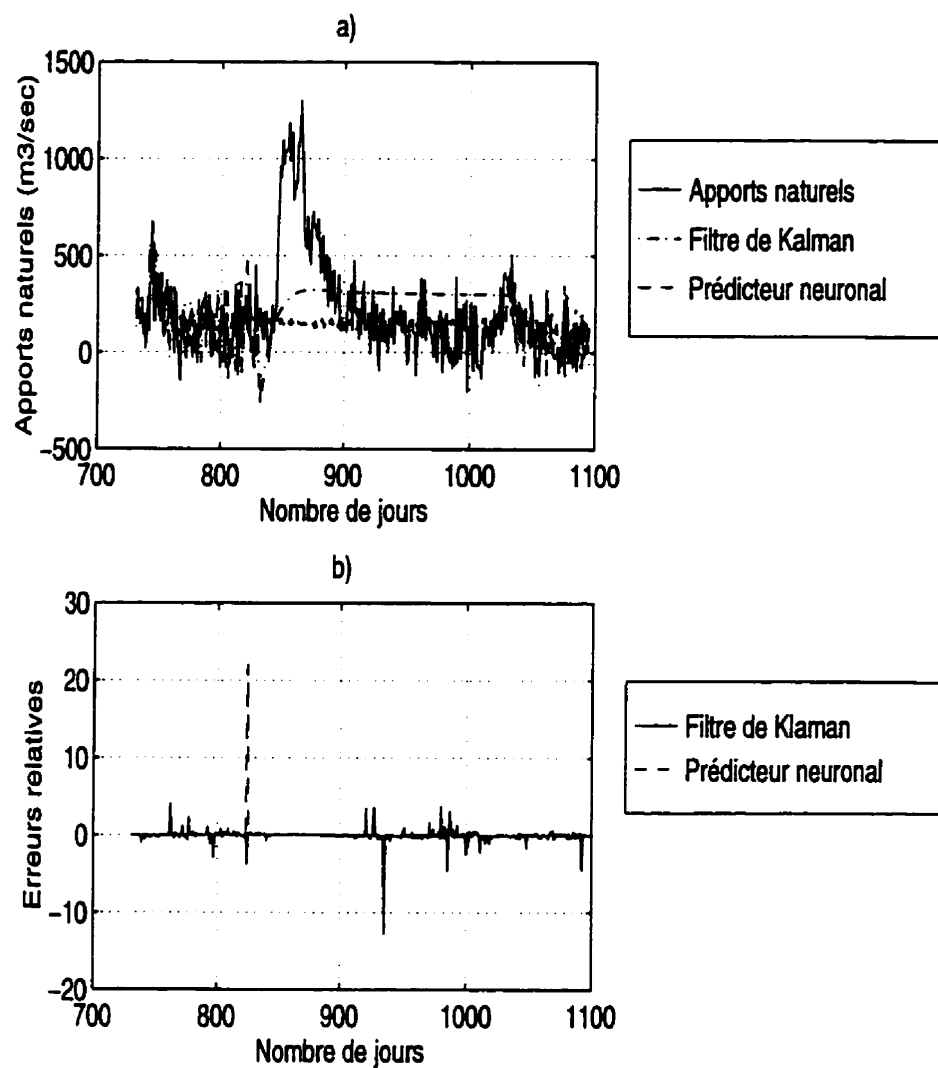
### **7.2.1 Le rapport bruit/signal à 5%**

Il faut noter ici, au regard des courbes des figures 7.13 ;7.14; 7.15, que le prédicteur neuronal ne donne plus beaucoup de bons estimés tel que le montre aussi le tableau 7.4. L'intensité du bruit des états à prédire a un impact négatif sur la bonne prédiction du filtre neuronal.

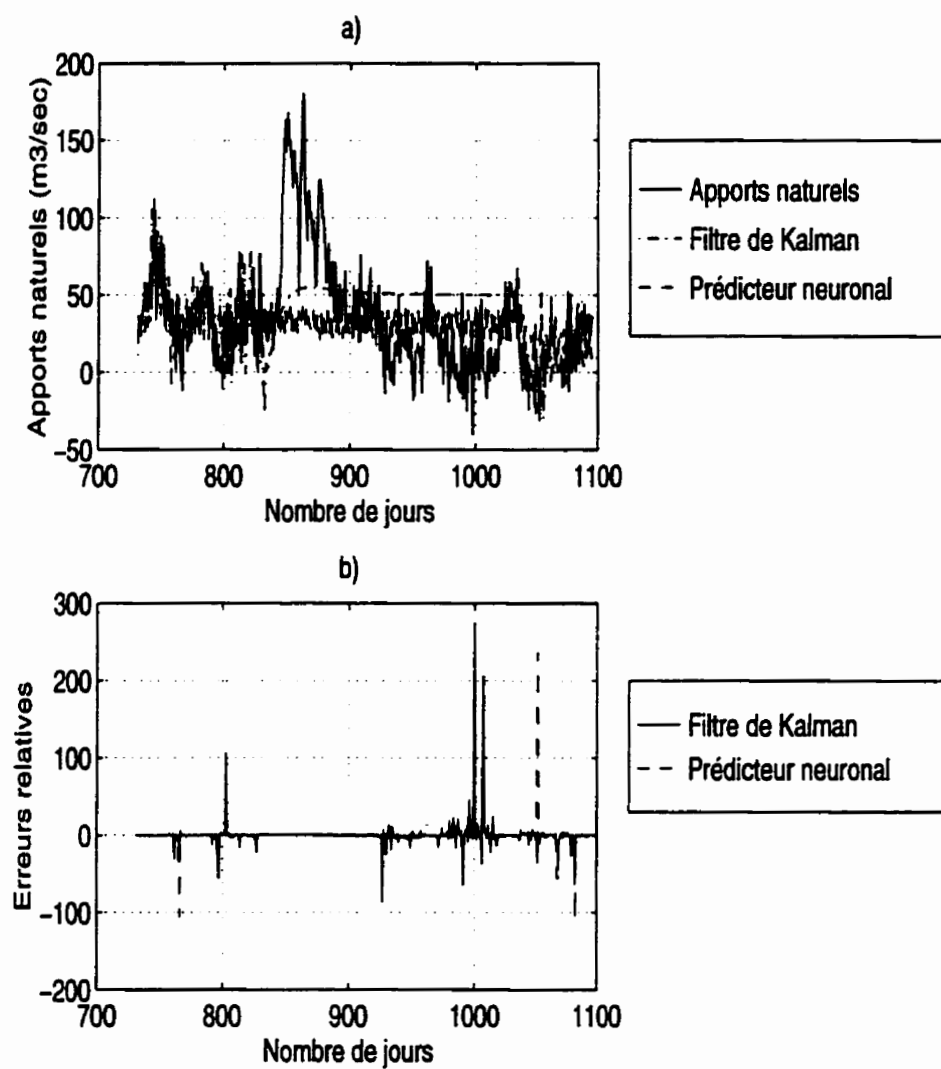
**Tableau 7.4:** Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures pour un rapport bruit/signal à 5%

Jours	Cabonga			Maniwaki			Baskatong		
	Kalman	Reseau	Reel	Reseau	Kalman	Reel	Kalman	Reseau	Reel
831	-9.16	31.43	28.76	16.41	-7.43	15.93	-118.09	153.02	147.00
832	-27.23	24.39	24.90	14.01	-22.23	23.36	-283.42	125.79	229.0
833	5.80	28.40	21.43	13.47	-6.27	13.81	-106.63	119.57	122.9
834	-1.03	22.58	39.42	11.54	-13.00	10.65	-181.166	98.36	87.6
835	5.26	26.09	24.72	9.52	-9.19	21.28	-138.59	75.61	206.1
836	12.22	30.06	45.17	14.67	-3.05	11.29	-70.15	133.18	95.2
837	12.95	19.27	28.46	12.73	-0.27	21.65	-38.84	111.81	210.4
838	16.78	39.69	16.58	16.41	3.62	8.85	4.85	153.02	69.0
839	21.07	28.10	30.72	17.83	6.60	-0.004	38.12	168.75	-29.9
840	24.76	34.09	40.54	16.10	8.31	4.63	57.37	149.25	22.8
841	26.90	39.60	54.90	15.56	10.11	12.18	77.45	143.84	105.9
842	32.35	25.38	34.07	17.84	11.88	24.79	97.27	168.91	245.9
843	34.98	40.04	53.63	15.96	13.76	11.43	118.32	148.09	98.7
844	37.17	28.66	68.55	16.02	15.80	19.27	141.15	170.84	187.2
845	39.03	34.21	117.00	16.50	17.71	22.70	162.64	153.74	225.8
846	40.80	39.79	141.79	15.67	19.37	50.73	181.23	144.83	559.3
847	42.60	25.34	163.97	17.84	20.74	70.80	196.60	168.85	764.2
848	44.36	40.01	142.72	15.93	21.87	88.64	209.23	147.75	965.2
849	45.97	28.65	168.36	17.99	22.86	83.03	220.39	170.60	904.9
850	47.36	19.27	167.45	16.49	23.81	100.78	231.06	153.63	1100.6

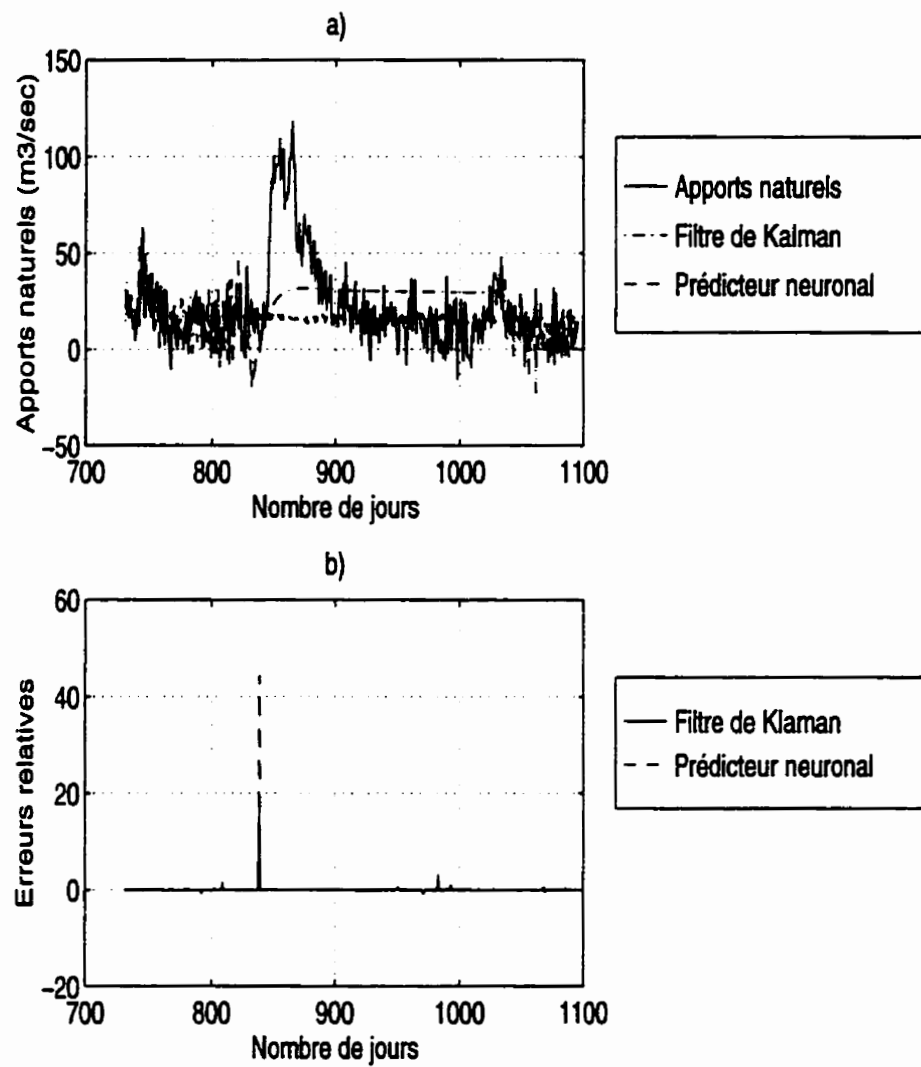




**Figure 7.13:** a) Prédiction à observations incomplètes rapport bruit/signal à 5%; b) Erreurs relatives (Baskatong).



**Figure 7.14:** a) Prédiction à observations incomplètes rapport bruit/signal à 5%; b) Erreurs relatives (Cabonga).



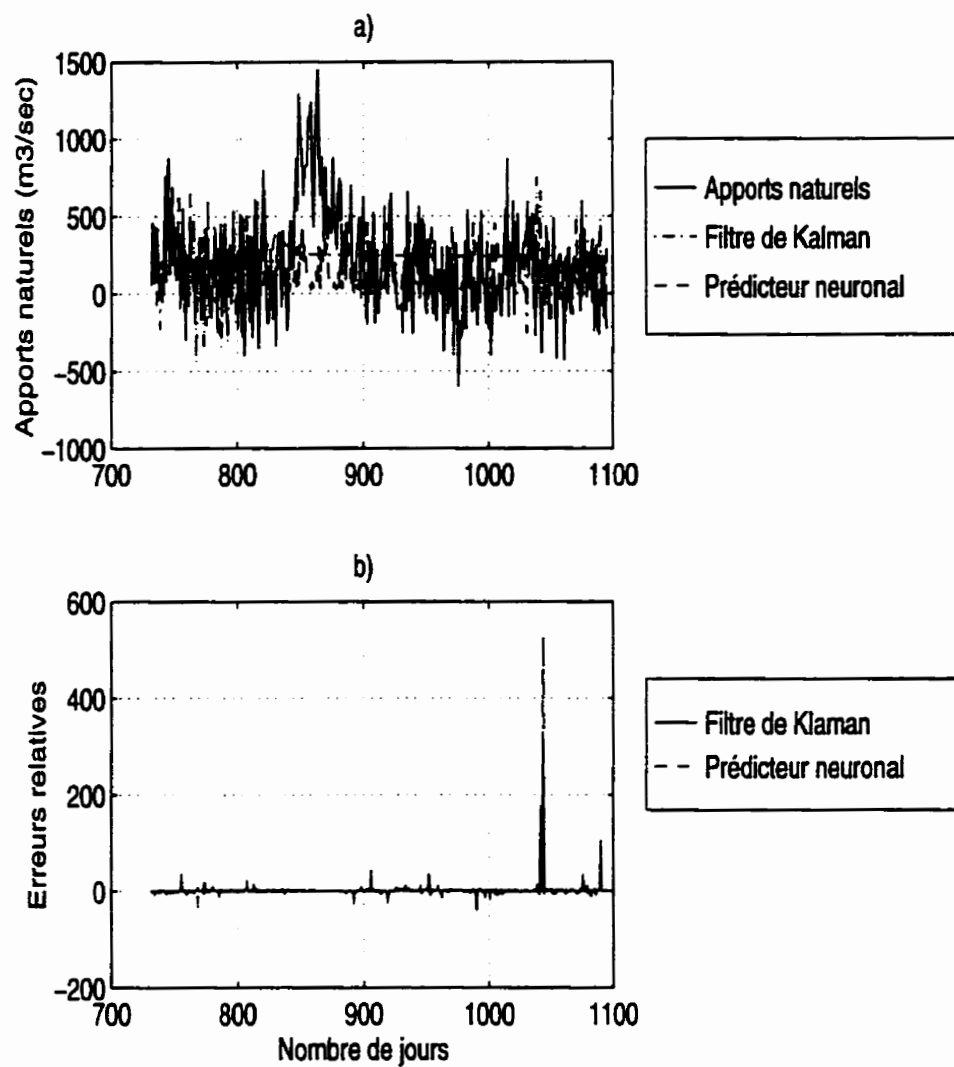
**Figure 7.15:** a) Prédiction à observations incomplètes rapport signal-bruit à 5%; b) Erreurs relatives (Maniwaki).

### 7.2.2 Le rapport bruit/signal à 10%

Comme dans le cas de la prédiction à mesures complètes, on remarque (7.16; 7.17; 7.18;) que les deux filtres ne peuvent plus faire la bonne prédiction lorsque les données sont incomplètes. Le tableau 7.5 montre les mesures prédites obtenues par le filtre de Kalman et le prédicteur neuronal entre la 831<sup>eme</sup> et la 850<sup>eme</sup> mesure.

**Tableau 7.5:** Données prédites par le filtre de Kalman et le prédicteur neuronal en l'absence des mesures pour un rapport bruit/signal à 10%

Jours	Cabonga			Maniwaki			Baskatong		
	Kalman	Reseau	Reel	Kalman	Reseau	Reel	Kalman	Reseau	Reel
831	62.60	33.12	-43.17	37.45	17.88	-28.16	387.78	169.77	-341.9
832	64.89	37.20	41.63	40.40	24.45	23.98	420.79	247.88	235.9
833	59.02	33.12	60.99	37.77	17.88	35.94	391.69	169.77	368.3
834	57.86	38.82	64.90	37.16	19.66	37.31	384.88	189.84	383.2
835	56.34	62.35	35.37	35.67	38.62	18.79	368.26	400.42	178.5
836	54.94	33.12	-1.95	34.07	17.88	-5.06	350.45	169.77	-86.2
837	53.93	33.60	74.65	32.95	26.70	39.73	337.91	284.97	410.9
838	52.48	18.67	25.99	31.96	10.69	7.34	326.88	85.62	52.3
839	50.92	15.05	46.92	31.21	9.17	17.36	318.54	68.26	162.8
840	49.45	15.04	-7.88	30.59	9.16	-19.03	311.72	68.19	-239.8
841	48.21	15.04	5.63	30.00	9.16	-9.21	305.10	68.19	-131.5
842	47.32	3.94	73.98	29.41	5.55	36.49	298.56	29.75	375.7
843	46.68	15.04	89.67	28.83	9.16	45.53	292.15	68.19	476.8
844	46.19	3.94	49.77	28.31	5.55	16.90	286.29	29.75	161.0
845	45.72	3.94	60.35	27.86	5.55	17.66	281.31	29.75	170.0
846	45.23	15.10	127.03	27.50	16.10	56.88	277.28	148.12	607.5
847	44.73	3.94	158.63	27.21	5.55	61.12	274.02	29.75	878.7
848	44.27	26.96	100.98	26.96	30.02	50.01	271.24	305.49	536.8
849	43.89	3.94	201.02	26.73	5.55	118.77	268.70	29.75	1301.3
850	43.60	40.44	174.35	26.51	30.42	104.46	266.28	315.64	1141.3



**Figure 7.16:** a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Baskatong).

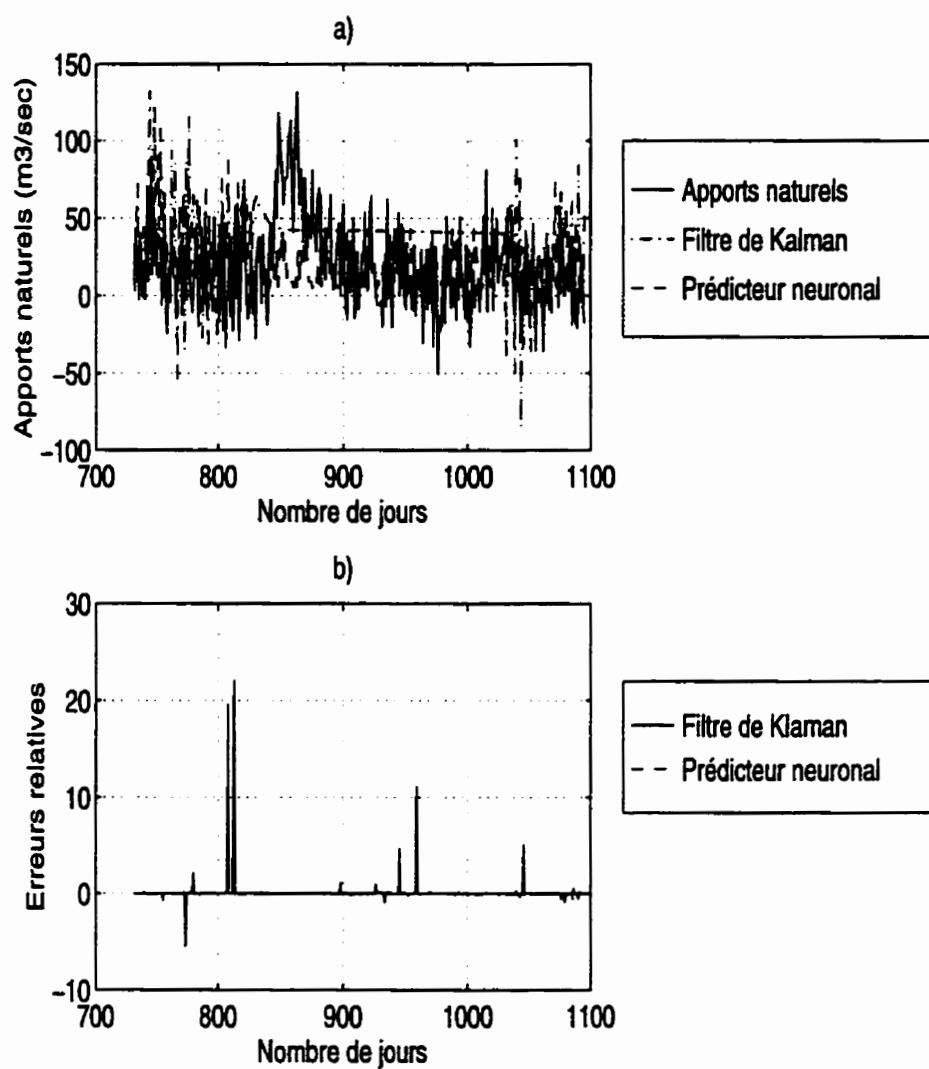
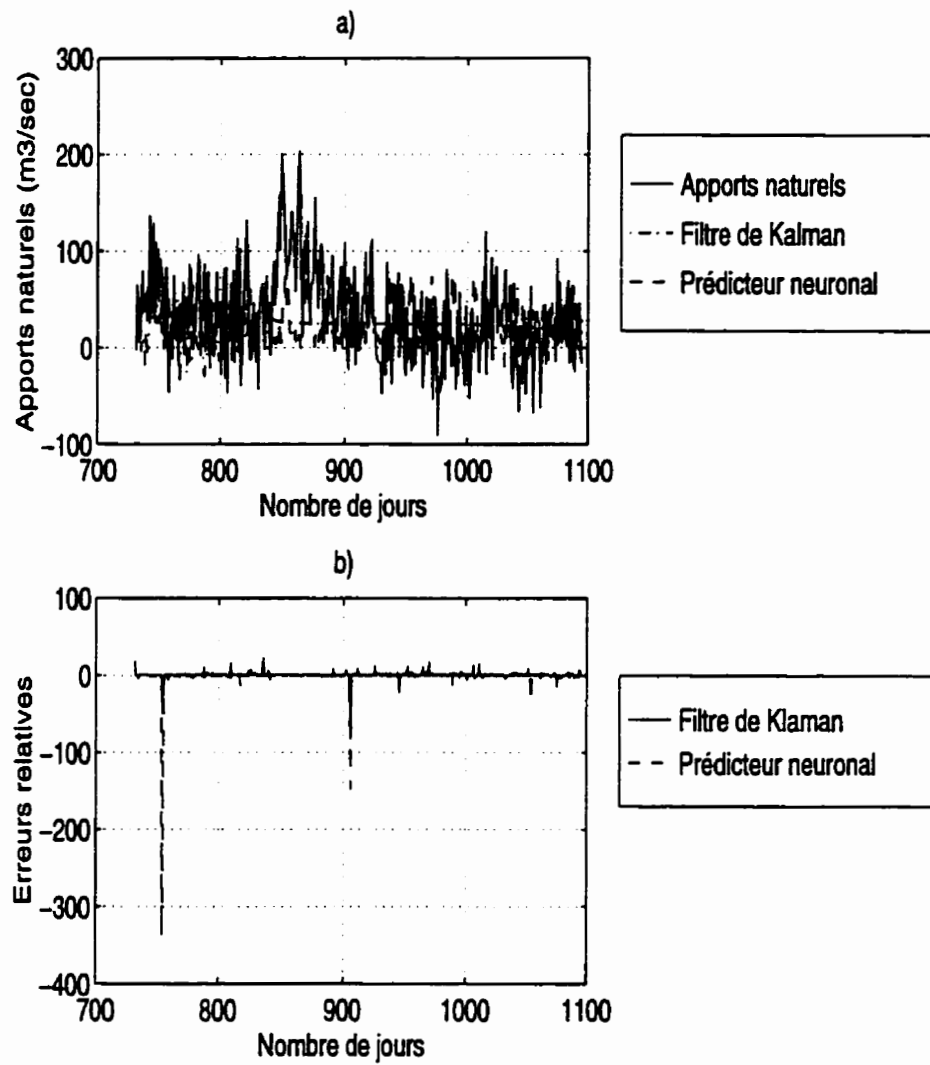


Figure 7.17: a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Maniwaki).



**Figure 7.18:** a) Prédiction à observations incomplètes rapport bruit/signal à 10%; b) Erreurs relatives (Cabonga).



## Conclusion

L'étude comparative des résultats donnés par les deux filtres; le filtre de Kalman et le prédicteur neuronal, dans un environnement stochastique a montré qu'ils ont un comportement semblable en prédiction. De même l'on observe que la qualité des résultats du prédicteur neuronal est fonction du degré de bruit du système dont il faut prédire les états. Plus le système est corrompu, plus le prédicteur neuronal détériore sa performance. Les simulations à 5% et à 10% (rapport bruit/signal) précédents ont validé ce constat. L'utilisation des mesures passées comme entrées au réseau permet d'améliorer la capacité de mémorisation du prédicteur. Cette structure s'écarte un peu de celle de Kalman; mais il n'en demeure pas moins vrai que les deux prédicteurs soient fondamentalement basés sur la même structure.

## CONCLUSION GÉNÉRALE

L'objet de ce travail a été de pouvoir analyser le design d'un filtre neuronal basé sur la structure du filtre de Kalman afin d'estimer et de prédire les états des systèmes dynamiques dans un environnement stochastique où, d'une part, les statistiques de leurs bruits sont inconnues et d'autre part, leur modèle est perturbé. L'objectif a été atteint.

Après une brève étude des méthodes utilisées à l'heure actuelle pour estimer le gain du filtre de Kalman, nous avons constaté que ces méthodes ne s'appliquent que dans le cadre très limitatif des bruits blancs gaussiens. Il fallait à cet effet imaginer un estimateur qui ne tiendrait pas compte des statistiques des bruits de modèle et de mesure et qui serait aussi quasi insensible aux variations des paramètres.

Le chapitre 4 a présenté la structure de cet estimateur qui remplace le gain du filtre de Kalman par un réseau de neurones artificiels. Les résultats de simulation obtenus, en considérant l'estimation des états du mouvement longitudinal de l'avion Boeing 747, ont validé ce nouvel estimateur. Deux types de bruits ont été considérés; des bruits blancs gaussiens non corrélés et des bruits non blancs et non gaussiens avec des statistiques inconnues. L'hypothèse des bruits blancs gaussiens a été prise en compte dans le souci de procéder à une comparaison entre l'estimateur neuronal et le filtre de Kalman pour lequel les statistiques des bruits sont nécessaires à son implantation. Il ressort de cet examen des rapports de ressemblance et de différence, que le filtre de Kalman donne de mauvais estimés si la matrice d'état

du système est perturbée. En modifiant la matrice d'état  $F$  par variation des pôles en boucle ouverte de 5% et en posant la matrice des entrées  $B$  ( $B_1 = 2 * B$ ) on constate effectivement que le filtre de Kalman ne donne plus de bons estimés pour les quatre états du système. Ceci fait remarquer que dans l'utilisation du filtre de Kalman, non seulement les statistiques des bruits du système doivent être connues, mais qu'il faut aussi supposer que les matrices d'état et les matrices des entrées ne varient pas dans le temps. Ce comportement est prévisible par l'équation (6.17) qui montre l'erreur d'estimation comme fonction de la variation des matrices du système perturbé. L'estimateur neuronal, dans ce contexte, est quasi insensible à la variation des paramètres du système et donne des résultats acceptables.

La non connaissance des statistiques des bruits du système rend impossible l'implantation du filtre de Kalman. Lorsque les bruits sont non blancs et non gaussiens (quelconques), le filtre de Kalman ne peut plus estimer les états. L'estimateur neuronal au contraire, arrive à faire une assez bonne estimation.

L'utilisation du filtre neuronal comme prédicteur, a permis de calculer les prédictions à  $m$  pas ( $m > 1$ ) des états des systèmes dynamiques. Les résultats de simulation montrent qu'il est possible de prédire véritablement une dizaine de valeurs des apports naturels de trois réservoirs (Baskatong, Cabonga, Maniwaki) d'Hydro-Québec. La sortie du filtre de Kalman tend vers une valeur constante dès que les informations de sortie deviennent inconnues si le système a un intégrateur; et tend vers zéro si le système est stable. Cependant, l'analyse de l'estimateur neuronal et du prédicteur dans le cas des systèmes suffisamment corrompus montre que le filtre neuronal a des limites. Si le rapport bruit/signal est supérieur à 5% les résultats de simulation prouvent que le filtre neuronal devient inopérant.

Cette thèse est un outil de conception du filtre neuronal et elle en fait une com-

paraïson avec le filtre de Kalman. Les résultats obtenus montrent bien qu'en cas d'absence des statistiques des bruits de modèle et de mesure, comme nous l'avons déjà précisé, le filtre neuronal fait une bonne estimation et une bonne prédiction des états d'un système moins corrompu même si les paramètres du système sont mal connus.

La comparaison que nous avons faite entre le filtre de Kalman et le filtre neuronal, dans le cas de la prédiction, ne tient pas compte du temps d'apprentissage des réseaux. L'apprentissage est une étape importante dans le processus de prédiction par le filtre neuronal. Par exemple dans le cadre de cette recherche, pour couvrir 12000 *epoch*, la durée moyenne d'apprentissage a été de huit heures sur UNIX SPARCstation 5 (Wiener). C'est en mode de rappel que le filtre neuronal prédit les apports naturels des réservoirs. Nous avons donc comparé ses estimés avec ceux donnés par le filtre de Kalman.

Si les statistiques des bruits inhérents au système et son modèle sont parfaitement connus, l'utilisation du filtre neuronal n'est pas nécessaire. Le filtre de Kalman donne dans ce cas des estimés optimaux avec un temps de calcul relativement court. Ainsi si les bruits sont blancs et gaussiens, le filtre de Kalman reste le meilleur estimateur et il ne sera pas avantageux d'utiliser le filtre neuronal. Par contre, dans le contexte de bruits inconnus et de modèle incertain, le filtre neuronal est l'estimateur idéal qu'il faudra utiliser bien qu'il soit moins performant si le rapport bruit/signal est considérable. En effet les résultats de simulation du filtre neuronal en prédiction et en estimation ont montré une dépendance entre le degré de bruit du système dont on veut prédire ou estimer les états, et les performances du filtre. Si le système est très corrompu, les résultats produits par le filtre se détériorent. En somme, le filtre neuronal joue un rôle complémentaire au filtre de Kalman dont les critères de

fonctionnement et de convergence sont clairement connus. Si ceux-ci ne sont pas remplis, il n'est pas indiqué de l'utiliser.

L'étude du comportement et du fonctionnement des réseaux neuronaux artificiels, dans un cadre mathématique rigoureux, reste beaucoup à faire. Aussi la caractérisation rigoureuse du filtre neuronal prendra davantage forme lorsque les réseaux neuronaux bénéficieront d'un formalisme mathématique solide. Cela devrait faciliter, d'une façon importante, la synthèse du filtre neuronal pour la prédiction et l'estimation des états dans un environnement stochastique.

Le filtre neuronal associé au procédé dont il estime les états constituent un système dynamique. L'analyse de sa stabilité est donc nécessaire afin de circonscrire les zones de convergence des estimés comme c'est le cas dans la littérature pour le filtre de Kalman. Le filtre neuronal est un système non linéaire à cause de la structure des réseaux neuronaux et leurs fonctions d'activation. L'étude de sa stabilité doit donc faire appel aux méthodes d'analyse non linéaires. Dans un travail futur, il serait intéressant de voir dans quelle mesure on pourrait utiliser quelques théorèmes tels que le théorème de la passivité, ou le théorème de la condition de secteur, ou même celui de la ligne de Popov et du petit gain pour trouver les conditions de stabilité du filtre neuronal.

## BIBLIOGRAPHIE

**ASSAF, H.** et **QUICK, M. C.** (1991), Updating hydrological forecasts, Revue Canadienne de Génie Civil, vol.18, 663-674.

**BAVARIAN, B.** (1988), Introduction to neural networks for intelligent control, IEEE Control Systems Magazine, 3-7.

**BERGMAN, M. J.** et **DELLEUR J. W.** (1985), Kalman filter estimation and prediction of daily stream flows: I. review, algorithm and simulation experiments, Water Resources Bulletin, vol.21, No.5, 815-825.

**BRAULT, J.J.** (1994), Cours gradué ELE6307 machines neuronales: architecture et applications, Ecole Polytechnique, DGEI, Montréal.

**BRYSON, A. E.** (1994), Control of spacecraft and aircraft, Princeton University Press, Princeton, New Jersey.

**BULSARI, A.** (1993), Multivariable control of linear systems using feedforward neural networks, Second IEEE Conf. on Control Applications, September 13-16, Vancouver, B. C. 319-327.

**CARPENTER, G. A. et GROSSBERG, S. (1987)**, ART2: Self-Organisation of stable category recognition codes for analog input patterns, Applied Optics, 26(23), 4919-4930.

**CHEN, F. et CHUANG, N. (1990)**, Back-Propagation neural networks for non-linear self-tuning adaptive control, IEEE Control systems Magazine, 44-48.

**DAVIS, M. H. A. et VARAIYA, P. (1973)**, Dynamic programming conditions for partially observable stochastic systems, SIAM Journal of Control, vol.11, No2, 226-232.

**DEMUTH, H. et BEALE, M. (1994)**, Neural network toolbox for use with matlab, The Math Work Toolbox Inc.

**FLETCHER, R. et POWELL, F. (1963)**, A rapidly convergence descent method for minimization, Comput. J., vol.7, No 2, 163-168.

**FLETCHER, R. (1987)**, Practical method of minimization, John Wiley, New York.

**GABOR, D. (1954)**, Communication theory and cybernetics, IRE Trans. on Circuit Theory, CT-1 19-31.

**GROSSBERG, S. (1982)**, Studies of mind and brain: neural principles of learning, perception, development, cognition and motor control, Reidel Press, Boston.

**GUEZ, A., EILBERT, J.L. et MOSHE, K. (1987)**, Neuromorphic architecture adaptive control: a preliminary analysis, In Proc. IEEE int. conf. neural networks, San Diego, CA vol.4, 567-572.

**HEBB, D. (1949)**, Organization of Behavior, John Wiley and Sons, New York.

**HECHT-NIELSEN, R. (1990)**, Neurocomputing, Addison-Wesley.

**HOPFIELD, J.J. (1982)**, Neural networks and physical systems with emergent collective computational abilities, Proceedings of the national academy of sciences, 79, 2554-2558.

**HIGUNI, Y., HIDEAKI, S. et HIDEKATSU T. (1991)**, A nonlinear regulator design in the presence of system uncertainties using multilayered neural networks, IEEE Trans. on Neural Networks, vol.2, No.4, 410-417.

**JACOBS, R. A. (1988)**, Increased rates of convergence through learning rate adaptation, Neural Networks, vol.1, pp 295-307.

**JER-NAN, J. CHUNG-WEN, C. et Minh, P. (1993)**, Estimation of Kalman filter gain from output residuals, Journal of Guidance, Control and Dynamics, vol.16, No.5, 903-908.

**KALMAN, R. E. (1960)**, A new approach to linear filtering and prediction prob-



lems", Trans. ASME, J. Basic Engrg., Vol.82, 34-45.

**KANDIL, N., KHORASANI, K., PATEL, R. V. et SOOD, V. K.** (1993), Optimum learning rate for backpropagation neural network, Canadian Conference on Electrical and Computer Engineering, Vancouver, Sept., 465-468.

**KIMBLE, G. A.** (1961), Hilgard and Maquis' conditioning and learning, Appleton-Century-Crofts, New York.

**KOJI, T.**, (1978), Estimation of steady-state Kalman filter gain, IEEE Trans. on Automatic Control, vol.Ac-23, No.5, 944-945.

**KOSKO, B.** (1988), Bidirectional association memories, IEEE Trans. Systems, Man and Cyber., 18(1), 49-60.

**KOHONEN, T.** (1988), Self-Organization and associative memories, Second Edition 1(1), 3, 3-16.

**KUMAR, P.R. et VARAIYA, P.** (1986), Stochastic systems, estimation identification and adaptive control, Prentice-Hall Information and Systems Series, New Jersey.

**KUMAR, S. et GUEZ, A.** (1991), ART based adaptive pole placement for neurocontrollers, Neural Networks, vol.4, 319-335.

**KUSCHEWKI, J.G., HUI T. et ZAK, S. H. (1993)**, Application of feedforward neural networks to dynamical systems identification and control, IEEE Trans. on Control Systems Technology, vol.1, No.1, 37-49.

**LAPEDES, A. et FARBER, R. (1987)**, Nonlinear signal processing using neural networks: prediction and system modeling, Technical Report, LA-UR-\*8-2662, Los Alamos National Laboratory.

**LEVIN, E. et INBAR, R. G. F. (1991)**, Neural networks architecture for adaptive systems modeling and control, Neural Networks, vol.4, 185-191.

**LEWIS, F. L. (1986)**, Optimization with an introduction to stochastic control theory, John Wiley and Sons, 13.

**MALHAMÉ, P. R. (1994)**, Cours gradué ELE 6215 commande stochastique, Ecole Polytechnique DGEI Montréal.

**MALSBURG, C. (1973)**, Self-Organization of orientation sensitive cells in the striate cortex, Kybernetik, 14, 85-100.

**MARKUS, M. (1995)**, Application of neural networks in stream flow forecasting, Thèse de doctorat, Colorado University.

**MAYERS, A. et TAPKEY, H. (1976)**, Adaptive sequential estimation with

unknown noise statistics, IEEE Trans. Auto. Contr., vol.24, No 4, 520-523.

**McCULLOCH, W. et PITTS, W. (1943)**, A logical calculus of ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, 7, 115-133.

**MEHROTRA, K., MOHAN, C. K. et RANKA, S. (1997)**, Elements of artificial neural networks, The MIT Press, Cambridge, Massachusettes, London, England.

**MINSKY, M. (1954)**, Neural Nets and the brain: model problem, Thèse de doctorat, Princeton University, Princeton NJ.

**NAIM, P. (1990)**, Des Réseaux de neurones, Eyrolles.

**NARENDRA, K.S. et PARTHASARATHY, K. (1990)**, Identification and control of dynamical systems using neural networks, IEEE trans. on neural networks, vol.1, No.1, 4-27.

**NARENDRA, K. S. et PARTHASARATHY, K. (1991)**, Gradient methods for optimization of dynamical systems containing neural networks, IEEE Trans. on Neural Networks, vol.2, No.2, 252-262.

**NARENDRA, K. S. et MUKHOPADHYAY, S. (1994)**, Adaptive control of nonlinear multivariable systems using neural networks, Neural Networks, vol.7, No.5, 737-752.

NGUYEN, D. H. et WIDROW, B. (1990), Neural networks for self-learning control systems, IEEE Control Systems Magazine, vol.10, No.3, 18-23.

PSALTIS, D., SIDERIS, A. et YAMAMURA, A., A. (1987), A Multi-layered neural networks controller, IEEE Int. Conf. on Neural Networks, San Diego, 17-21.

PUSKORIUS, G.V.et FELDKAMP, L.A. (1994), Neurocontrol of nonlinear dynamical systems with Kalman filter-trained recurrent networks, IEEE Trans. on Neural Networks, vol.5, No.2, 279-297.

RAOL, J. R. (1995), Parameters estimation of state space models by recurrent neural networks, IEE Proc. Control Theory Appl., vol.142, No.2, 114-118.

ROBERTS, J.M., MILLS, D.J., CHARNLEY, D. et HARRIS, C.J. (1997), Improved Kalman filter initialisation using neurofuzzy estimation, [www.ecs.soton.ac.uk/publications/rj95/isis/jmr9ur/iee95.htm](http://www.ecs.soton.ac.uk/publications/rj95/isis/jmr9ur/iee95.htm).

ROSENBLATT, F. (1958), Perception: a probabilistic model for information storage and organization in the brain, Psychol. Rev., 65, 366-408.

ROSENBLATT, F. (1961), Principles of neurodynamics: perceptrons and the theory of brain mechanism, Spartan, New York.

RUCHTI, T. L., BRAWN, R.H et GARSIDE, J.J. (1992), A parameter estimation approach to artificial neural network weight selection for nonlinear system identification, IEEE Conference on Control Applications, vol.1, 565-570.

RUCK,D.W., ROGERS,S.K., MAYBECK,P.S. et ONLEY,M.E.(1992), Comparative analysis of back-propagation and the extended Kalman filter for training multi-layer perceptrons, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.14, No.6, 686-691.

SAAD, M., TURGEON, A., BIGRAS, P. et DUQUETTE, R. (1994), Learning disagregation technique for the operation of long-term hydroelectric power systems, Water Resources Research, vol.30, No.11, 3195-3202.

SHAH S., PALMIERI, F. et DATUM, M. (1992), Optimal filtering algorithm for fast learning in feedforward neural networks, Neural Networks, vol.5, 779-787.

SINGHAL, S. et WU, L. (1989), Training feedforward networks with the extended Kalman algorithm, Proc. IEEE int. conf. acoust. speech, signal processing, 1187-1190.

SWEIGEND, A. et GERSHENFELD, N.A. (1993), Time series prediction: forecasting the future and understanding the past, Addison-Wesley, Reading, MA.

**SWINGLER, K. (1996),** Applying neural networks: a practical guide, Academic Press, San Diego.

**TODINI, E. (1978),** Mutually interactive state-parameter (MISP) estimation, Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources; Editor: C.L. Chiu. Proceedings of the American Geophysical Union Chapman Conference University of Pittsburgh, Pittsburgh, 135-151.

**XIANZHONG, C. et Kang G. S. (1993),** Direct control and coordination using neural networks, IEEE trans. on Systems, Man, and Cybernetics, vol.23, No.3, 686-697.

**XU, H. Y. et VUKOVICH, G. (1993),** An experimental study of robotic modeling and control using neural networks, Control Eng. Practice, vol.1, No. 4, 715-724.

**YOSHIAKI, I. et SAWA, T. (1992),** Neural networks application for direct feedback controllers, IEEE Trans. on Neural Networks, vol.3, No.2, 224-231.

# ANNEXE I

## Démonstration du théorème de l'existence du régime permanent du filtre de Kalman (comme rapportée par Malhamé, 1994)

### Procédure de démonstration

Ce théorème se démontre sur quatre étapes.

- Il faut montrer que pour  $\Sigma_{1|0} = 0$ , la covariance du prédictor  $\Sigma_{k|k-1}$  converge à l'état d'équilibre vers une matrice  $\bar{\Sigma}$  semi-définie positive satisfaisant l'équation de Riccati.
- Si  $\bar{L} = \bar{\Sigma}C^t[C\bar{\Sigma}C^t + R]^{-1}$  et  $\bar{K} = F\bar{L}$  (gain du prédictor de Kalman en régime permanent); alors  $(F - \bar{K}C)$  est une matrice strictement stable.
- Montrer alors que pour tout  $\Sigma_{1|0}$  de la forme  $\delta I$ ,  $\delta > 0$   $\Sigma_{k|k-1} \rightarrow \bar{\Sigma}$  et de plus, la solution de Riccati est unique parmi les matrices semi-définies positives.
- Généraliser le résultat précédent pour  $\Sigma_{1|0}$  arbitraire, semi-définie positive

### Étape 1

Soit  $\Sigma_{k|k-1}$  correspondant à  $\Sigma_{0|-1} = 0$ , et  $\bar{\Sigma}_{k|k-1}$  correspondant à  $\bar{\Sigma}_{0|-1} = 0$ , par le lemme de la monotonie on a  $\Sigma_{1|0} \geq 0$  et  $\Sigma_{0|-1} = 0$  donc  $\Sigma_{1|0} \geq \Sigma_{0|-1}$  ce qui nous permet de déduire que  $\Sigma_{k+1|k} \geq \bar{\Sigma}_{k|k-1}$  par ailleurs  $\Sigma_{k+1|k} \geq \Sigma_{k|k-1}$ .

La matrice  $\Sigma_{k|k-1}$  est croissante mais également bornée par en dessus:  $\Sigma_{k|k-1} \leq \Sigma_{k+1|k} \leq \dots$ . En effet  $\forall k \Sigma_{k|k-1} \leq \psi$ .

Pour les termes diagonaux, la suite

$$e_j^t [\Sigma_{k|k-1}] e_j = \Sigma_{k|k-1}^{(j,j)} \quad (7.1)$$

où

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ i \\ 0 \\ \vdots \\ \vdots \end{pmatrix} \leftarrow j^{ieme} \text{ colonne} \quad (7.2)$$

est croissante puisque

$$e_j^t [\Sigma_{k+1|k} - \Sigma_{k|k-1}] e_j = \Sigma_{k+1|k}^{(j,j)} - \Sigma_{k|k-1}^{(j,j)} \geq 0 \quad (7.3)$$

Les termes diagonaux convergent vers une limite (unique). Pour les termes non-diagonaux, on utilise:

$$(e_i + e_j)^t [\Sigma_{k+1|k} - \Sigma_{k|k-1}] (e_i + e_j) \geq 0 \quad (7.4)$$

$$\Sigma_{k+1|k}^{(i,i)} - \Sigma_{k|k-1}^{(i,i)} + \Sigma_{k+1|k}^{(j,j)} - \Sigma_{k|k-1}^{(j,j)} + 2(\Sigma_{k+1|k}^{(i,j)} - \Sigma_{k|k-1}^{(i,j)}) \geq 0 \quad (7.5)$$

par ailleurs,

$$\Sigma_{k+1|k}^{(i,i)} + \Sigma_{k+1|k}^{(j,j)} + 2\Sigma_{k+1|k}^{(i,j)} \quad (7.6)$$



est croissante et bornée, donc elle converge lorsque  $k \rightarrow \infty$ . Mais comme  $\Sigma_{k+1|k}^{(i,i)}$  et  $\Sigma_{k+1|k}^{(j,j)}$  convergent individuellement alors  $\Sigma_{k+1|k}^{(i,j)}$  doit converger; ce qui permet de voir que  $\Sigma_{k+1|k}$  converge vers une limite  $\bar{\Sigma}$ . Il est clair que cette limite devra satisfaire l'équation de Riccati obtenue en exprimant dans le filtre de Kalman  $\Sigma_{k+1|k}$  en fonction de  $\Sigma_{k|k-1}$  et en remplaçant  $\Sigma_{k+1|k}$  et  $\Sigma_{k|k-1}$  par  $\bar{\Sigma}$ .

## Étape 2

Démontrons que  $(F - \bar{K}C)$  est stable.

On peut vérifier que l'équation algébrique de Riccati peut s'écrire sous la forme:

$$\bar{\Sigma} = (F - \bar{K}C)\bar{\Sigma}(F - \bar{K}C)^t + \bar{K}R\bar{K}^t + GQG^t \quad (7.7)$$

Montrons à partir de l'équation (7.7) que  $(F - \bar{K}C)$  doit être stable. En effet supposons que  $\lambda$  est une valeur propre de  $(F - \bar{K}C)$  avec  $|\lambda| \geq 1$ .

soit  $\nu$  le vecteur propre associé à gauche  $(F - \bar{K}C)^t \nu = \lambda \nu$  de l'équation (7.7) on peut écrire:

$$\nu^* \bar{\Sigma} \nu = \nu^* (F - \bar{K}C) \bar{\Sigma} (F - \bar{K}C)^t \nu + \nu^* \bar{K}R\bar{K}^t \nu + \nu^* GQG^t \nu \quad (7.8)$$

où  $\nu^*$  est le conjugué complexe transposé. De la relation (7.8) nous pouvons déduire celle ci-dessous:

$$\nu^* \bar{\Sigma} \nu (1 - |\lambda|^2) = \nu^* \bar{K}R\bar{K}^t \nu + \nu^* GQG^t \nu \quad (7.9)$$

Dans l'équation (7.9),  $\nu^* \bar{\Sigma} \nu (1 - |\lambda|^2) \leq 0$ ,  $\nu^* GQG^t \nu \geq 0$  et  $\nu^* \bar{K}R\bar{K}^t \nu \geq 0$ .

L'égalité ne peut persister que si les deux membres de l'équation sont nuls. Ceci qui nous permet de voir que  $\nu^* \bar{K}R\bar{K}^t \nu = 0$  et  $\nu^* GQG^t \nu = 0$

$\nu^* \bar{K}R\bar{K}^t \nu$  étant égale à zéro, comme  $R > 0$ , on voit que  $\bar{K}^t \nu = 0$ .

L'expression  $\nu^* GQG^t \nu$  étant égale aussi à zéro, on déduit que  $(G\sqrt{Q})^t \nu = 0$  ce qui

revient à poser que  $C^t \bar{K}^t \nu = 0$ ; mais  $F^t \nu - C^t \bar{K}^t \nu = \lambda \nu$ ;  $\nu$  est donc un vecteur propre de  $F^t$ . Le système ci-dessous:

$$\begin{cases} F^t \nu = \lambda \nu \\ (G\sqrt{Q})^t \nu = 0 \end{cases} \quad (7.10)$$

montre que la paire  $(F^t, (G\sqrt{Q})^t)$  est non observable; autrement dit  $(F, G\sqrt{Q})$  est non commandable ce qui est contraire à l'hypothèse; donc  $(F - \bar{K}C)$  est strictement stable.

### Étape 3

Démontrons que pour  $\Sigma_{0|-1}$  de la forme  $\delta I$ ,  $\delta > 0$ , on aura  $\Sigma_{k|k-1} \rightarrow \bar{\Sigma}$ .

On peut voir que:

$$\Sigma_{k+1|k} - \bar{\Sigma} = (F - \bar{K}C)^{k+1}(\Sigma_{0|-1} - \bar{\Sigma})\phi^t(k+1, 0) \quad (7.11)$$

avec:

$$\phi(k, j) = (F - FL_{k+1}C) \dots (F - FL_jC) \quad (7.12)$$

Pour  $k \rightarrow \infty$ ;  $(F - \bar{K}C)^{k+1} \rightarrow 0$  (parce que les valeurs propres de  $(F - \bar{K}C)^{k+1}$  sont inférieures à 1) donc  $\Sigma_{k|k-1} \rightarrow \bar{\Sigma}$  si  $\phi(k+1, 0)$  reste bornée. Or  $\phi(k+1, 0)$  est bornée puisque:

$$\begin{aligned} \Sigma_{k+1|k} &= (F - FL_kC)\Sigma_{k|k-1}(F - FL_kC)^t + FL_kRL_k^tF^t + GQG^t \geq \\ &\quad (F - FL_kC)\Sigma_{k|k-1}(F - FL_kC)^t \end{aligned} \quad (7.13)$$

et en répétant l'inégalité:

$$\Sigma_{k+1|k} \geq \delta \phi(k+1, 0)\phi^t(k+1, 0) \quad (7.14)$$

Comme  $\Sigma_{k+1|k}$  est bornée elle-même, alors  $\phi(k+1, 0)$  est bornée et

$$\lim_{k \rightarrow \infty} \Sigma_{k|k-1} \rightarrow \bar{\Sigma} \quad (7.15)$$

#### Étape 4

Pour  $\Sigma_{0|-1}$  arbitraire semi-définie positive

$$\Sigma_{0|-1} = T \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} T^{-1} \quad \lambda_i \geq 0 \text{ reelles} \quad (7.16)$$

Prenons  $\lambda_{\max} = \max \lambda_i$

$$0 \leq \Sigma_{0|-1} \leq \lambda_{\max} [I]$$

puisque

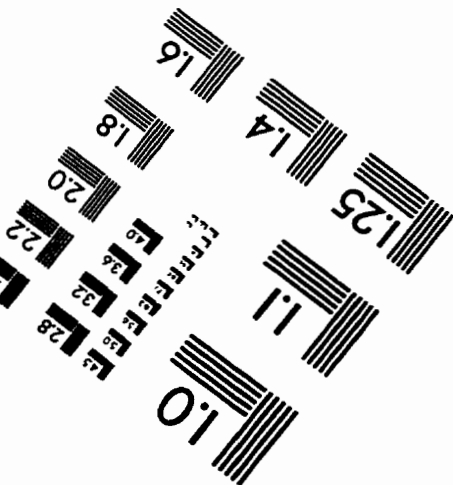
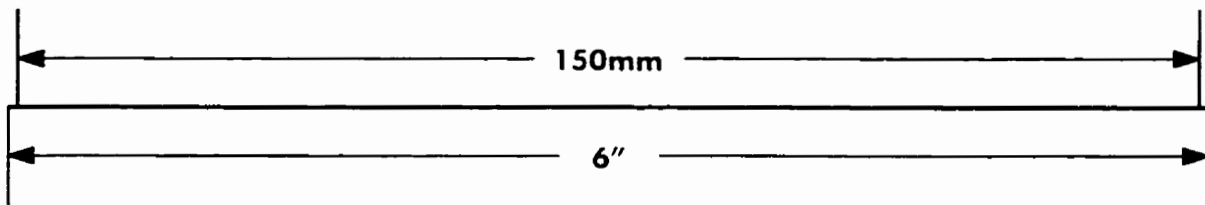
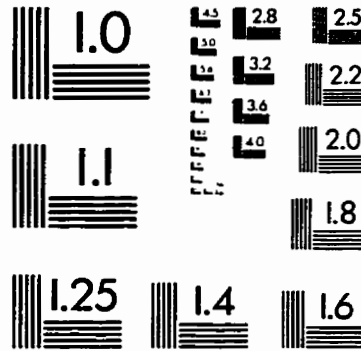
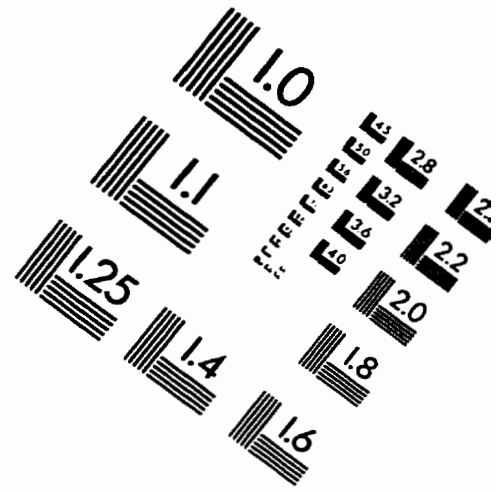
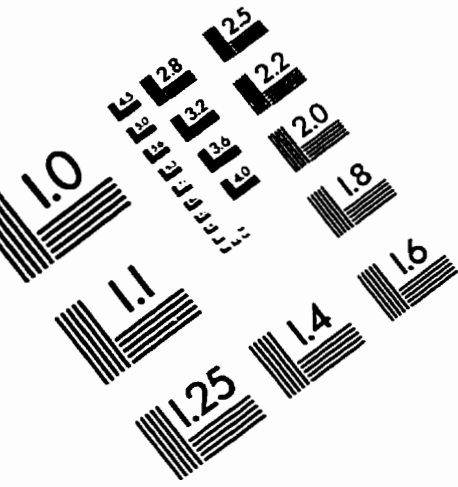
$$\Sigma_{0|-1} - \lambda_{\max} I = T \begin{bmatrix} \lambda_1 - \lambda_{\max} & & \\ & \ddots & \\ & & \lambda_n - \lambda_{\max} \end{bmatrix} T^{-1} \quad (7.17)$$

La solution de l'équation de Riccati est unique puisque si nous supposons  $\hat{\Sigma} \neq \bar{\Sigma}$  comme matrice de solution semi-définie positive de l'équation de Riccati, alors en posant  $\Sigma_{0|-1} = \hat{\Sigma}$ , il est clair que  $\Sigma_{k|k-1} = \hat{\Sigma} \quad \forall \quad k$  donc

$$\lim_{k \rightarrow \infty} \Sigma_{k|k-1} = \hat{\Sigma} \neq \bar{\Sigma} \quad (7.18)$$

ce qui est une contradiction.

# IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc  
1653 East Main Street  
Rochester, NY 14609 USA  
Phone: 716/482-0300  
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved

